Display
D

CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. X · NO. 2
1958

Schlicht Dirichlet series M. S. Robertson 161
Laplace transforms and generalized Laguerre polynomials P. G. Rooney 177
On some relations between partial and ordinary differential equations Erwin Kreyszig 183
A new proof of a theorem of Lehmer J. B. Roberts 191
Certain transformations of nearly-poised bilateral hypergeometric series of special type H. S. Shukla 195
Comma-free codes S. W. Golomb, Basil Gordon, and L. R. Welch 202
On the number of ordinary lines determined by n points L. M. Kelly and W. O. J. Moser 210
On the number of sides of a Petrie polygon Robert Steinberg 220
Central limit theorems for interchangeable processes J. R. Blum, H. Chernoff, M. Rosenblatt, and H. Teicher 222
Some generalizations of the problem of distinct representatives N. S. Mendelsohn and A. L. Dulmage 230
Algebraic approximation of curves A. H. Wallace 242
On certain pairs of matrices which generate free groups Bomshik Chang, S. A. Jennings and Rimhak Ree 279
A note on division algorithms in imaginary quadratic number fields D. W. Dubois and A. Steger 285
Dimension of ideals in polynomial rings Maurice Auslander and Alex Rosenberg 287
Groupoïdes automorphes par le groupe géométrique et quasigroupes "Endo" A. Sade 294

Published for
THE CANADIAN MATHEMATICAL CONGRESS
by the
University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, G. F. D. Duff, University of Toronto. Everything possible should be done to lighten the task of the reader; the notation and reference system should be carefully thought out. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is $8.00. This is reduced to $4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta
University of British Columbia
Dalhousie University
Université Laval
University of Manitoba
McMaster University
Queen's University
St. Mary's University

Assumption University
Carleton College
Ecole Polytechnique
Loyola College
McGill University
Université de Montréal
Royal Military College
University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

SCHLICHT DIRICHLET SERIES

M. S. ROBERTSON

1. Introduction. For power series

$$f(z) = z + a_2z^2 + \ldots + a_nz^n + \ldots \tag{1.1}$$

for which

$$\sum_2^\infty n|a_n| \leq 1, \tag{1.2}$$

it has been known for four decades **(1)** that $f(z)$ is regular and univalent or schlicht in $|z| < 1$. This theorem, due to J. W. Alexander, has more recently been studied by Remak **(5)** who has shown that $w = f(z)$, under the hypothesis (1.2), maps $|z| < 1$ onto a star-like region, and if (1.2) is not satisfied $f(z)$ need not be univalent in $|z| < 1$ for a proper choice of the amplitudes of the coefficients a_n.

We may recast the theorem of Alexander in the following form. Let the power series

$$f(z) = z + a_2z^2 + \ldots + a_nz^n + \ldots \tag{1.3}$$

have a radius of convergence $R > 0$, and let ρ be the largest positive number, $0 < \rho \leq R$, for which

$$\sum_2^\infty n|a_n|\rho^{n-1} \leq 1. \tag{1.4}$$

Then $f(z)$ is univalent and star-like with respect to the origin in $|z| < \rho$.

For Dirichlet series

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=2}^\infty a_n e^{-\lambda_n s}, \qquad s = \sigma + it, \tag{1.5}$$

whose abscissa of absolute convergence is $\bar{\sigma}$, $-\infty \leq \bar{\sigma} < \infty$, there is a smallest real number τ, $\bar{\sigma} \leq \tau < \infty$, for which

$$\sum_{n=2}^\infty \lambda_n|a_n|e^{-\lambda_n\tau} \leq 1. \tag{1.6}$$

Working by analogy with power series one might guess that under the hypothesis (1.6), $f(s)$ would be univalent in the half-plane $\Re s > \tau$. However, this is not the case, as the simple example

$$f(s) = -e^{-\lambda_1 s}$$

Received August 12, 1957. This paper was prepared at the Summer Research Institute of the Canadian Mathematical Congress.

shows, and because of the almost periodic character of the functions $f(s)$ in general.

Functions $f(z)$ given by power series (1.1) which satisfy (1.2) are said to be of Hurwitz class **(5)**. Similarly, those functions $f(s)$ given by Dirichlet series (1.5) which satisfy (1.6) will be said to be of class τ.

Recalling certain concepts of univalency introduced by Montel **(3)**, we say that $f(z)$ is *locally univalent* in a region D if $f(z)$ is regular in D and if, for every closed domain $D^* \subset D$ and for every point z_0 of D^*, there exists a positive number ρ independent of z_0 such that $f(z)$ is univalent in every circle $|z - z_0| < \rho$ lying within D. Moreover, if there is a class of functions $\{f(z)\}$ regular in the region D we shall say that the functions $f(z)$ of the class are *uniformly locally univalent* in D whenever $f(z)$ is locally univalent in D and ρ has the same value for each member $f(z)$ of the class.

We shall show that the functions $f(s)$ given by a Dirichlet series (1.5) of class τ are uniformly locally univalent in a half-plane. If $\tau < (\log \lambda_1)/\lambda_1$, the half-plane is the one for which $\Re s > \tau$. If $\tau \geqslant (\log \lambda_1)/\lambda_1$ the half-plane is the one for which

$$\Re s > \frac{\lambda_q \tau - \log \lambda_1}{\lambda_q - \lambda_1} \geqslant \tau,$$

where q is the suffix of the first non-vanishing coefficient a_q of the numbers a_n, $n \geqslant 2$. The theorem is a best possible one. More explicitly we prove

THEOREM 1. *Let*

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=q}^{\infty} a_n e^{-\lambda_n s}, a_q \neq 0, \qquad s = \sigma + it, \tag{1.7}$$

have $\bar{\sigma}$ as its abscissa of absolute convergence, $-\infty \leqslant \bar{\sigma} < \infty$. Let $f(s)$ be of class τ. Let ϵ be an arbitrary real number in the range $0 < \epsilon < 1$. Then $f(s)$ is univalent in every circle $|s - s_0| < (1 - \epsilon)\pi/\lambda_1$ for $\Re s_0 > \alpha$ where

$$\alpha = \max\left\{\tau + \frac{(1-\epsilon)\pi}{\lambda_1}, \frac{3 \log (2 - \epsilon) - \log (\lambda_1 \epsilon) + \lambda_q \tau + (1 - \epsilon)\pi\lambda_q/\lambda_1}{\lambda_q - \lambda_1}\right\}. \tag{1.8}$$

The factor π/λ_1 in the radius $(1 - \epsilon)\pi/\lambda_1$ cannot be replaced by a larger one.

An application is made to the Riemann Zeta-function $\zeta(s)$ which is shown to be locally schlicht in the half-plane $\Re s > 6.32$.

The radius of univalency of the function $e^{-\lambda_1 s}$ about any point s_0 is exactly π/λ_1 and the function has a period $2\pi i/\lambda_1$. Since this function is also univalent in every strip of width $2\pi/\lambda_1$ parallel to the real axis, this suggests that perhaps semi-infinite strips would form more natural domains in which to investigate properties of univalency for functions represented in half-planes by Dirichlet series. Accordingly, in §4 we obtain several results applicable to strip domains. The following theorem is proved.

THEOREM 3. *Let*

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=2}^{\infty} a_n e^{-\lambda_n s}, \qquad s = \sigma + it, \tag{1.9}$$

be absolutely convergent for $\sigma > \bar\sigma$, $-\infty < \bar\sigma < \infty$, and let $f(s)$ be of class $\tau < \tau_0$ where

$$\tau_0 = \frac{\log \lambda_1}{\lambda_1} - \frac{\log 2}{2\lambda_1} = \frac{\log \lambda_1 - \frac{1}{2}\log 2}{\lambda_1}. \tag{1.10}$$

Let k be an arbitrary integer and let

$$t_0 = \frac{1}{\lambda_1} \operatorname{arc\,cos}\left(\frac{e^{\lambda_1 \tau}}{\lambda_1}\right), \qquad 0 < t_0 < \frac{\pi}{2\lambda_1}. \tag{1.11}$$

Let D_k denote the strip of the s-plane defined by

$$\sigma > \tau, \quad \left| t - \frac{2k\pi}{\lambda_1} \right| < t_0.$$

Then $W = f(s)$ is univalent in D_k and maps the interior of D_k onto a bounded region Δ_k which is star-shaped with respect to the point $\sigma = +\infty$ at an end of the real axis, this region being convex in the direction of the real axis. The theorem is not true in general if the strip is enlarged, or if τ exceeds τ_0. If τ_0 is replaced by $\tau^*_0 = (\log \lambda_1)/\lambda_1$ $f(s)$ is still univalent in D_k, but Δ_k is in general no longer convex in the direction of the real axis. Again, the theorem is not true if τ exceeds τ^*_0.

2. Preliminary lemmas. Let $f(s)$ be defined by a Dirichlet series, and normalized as in (1.5), with $\bar\sigma$ as abscissa of absolute convergence, $-\infty < \bar\sigma < \infty$, and where λ_n is a given sequence

$$0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots, \lambda_n \to \infty.$$

We shall suppose that not all the coefficients a_n are zero.

It is well-known that the derived series

$$f'(s) = \lambda_1 e^{-\lambda_1 s} - \sum_{n=2}^{\infty} \lambda_n a_n e^{-\lambda_n s} \tag{2.1}$$

also converges absolutely for $\sigma > \bar\sigma$. If

$$\sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n \bar\sigma} \tag{2.2}$$

diverges to $+\infty$ and $\bar\sigma$ is finite there exists a unique real number τ, $\bar\sigma < \tau < \infty$, for which

$$\sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n \tau} = 1. \tag{2.3}$$

This follows since

$$g(\sigma) = \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n \sigma} \tag{2.4}$$

is a strictly decreasing continuous function of σ for $\sigma > \bar{\sigma}$ which assumes arbitrarily large positive values for σ near $\bar{\sigma}$, $\sigma > \bar{\sigma}$, and which assumes arbitrarily small positive values for sufficiently large positive values of σ. Since $\bar{\sigma}$ was assumed finite, there are an infinite number of coefficients a_n different from zero.

The same conclusion about τ in (2.3) may be drawn if the series (2.2) converges to a positive number not less than 1. In this case $\bar{\sigma} \leqslant \tau < \infty$. If the series (2.2) converges to a positive number less than 1 we define $\tau = \bar{\sigma}$ so that in this case (2.3) is replaced by

$$\sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n \tau} < 1. \tag{2.5}$$

If $\bar{\sigma} = -\infty$, $g(\sigma)$ assumes arbitrarily large positive values for τ sufficiently small (algebraically) and negative so that again there exists a unique τ for which (2.3) holds. In all cases $\bar{\sigma} \leqslant \tau < \infty$. We shall call τ the "class" of the Dirichlet series (1.5). Thus we have the lemma:

Lemma 1. *Let $f(s)$ be defined by the Dirichlet series* (1.5) *with $\bar{\sigma}$ as its abscissa of absolute convergence, $-\infty < \bar{\sigma} < \infty$. Then there exists a smallest real number, τ, $\bar{\sigma} \leqslant \tau < \infty$, for which*

$$\sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n \tau} \leqslant 1. \tag{2.6}$$

Lemma 2. *Let s_1 and s_2 be any two distinct points in the circle $|s| \leqslant r$. Let λ be a positive number. Then*

$$\left| \frac{e^{-\lambda s_2} - e^{-\lambda s_1}}{s_2 - s_1} \right| \leqslant \lambda e^{r\lambda}. \tag{2.7}$$

Lemma 2 follows immediately from the expansion

$$\begin{aligned} \left| \frac{e^{-\lambda s_2} - e^{-\lambda s_1}}{s_2 - s_1} \right| &= \lambda \left| 1 - \frac{(s_2 + s_1)}{2!} \lambda + \frac{(s_2^2 + s_2 s_1 + s_1^2)}{3!} \lambda^2 - \dots \right| \\ &\leqslant \lambda \left(1 + \frac{r\lambda}{1!} + \frac{r^2\lambda^2}{2!} + \dots + \frac{r^k \lambda^k}{k!} + \dots \right) = \lambda e^{r\lambda}. \end{aligned} \tag{2.8}$$

Lemma 3. *Let s_1 and s_2 be any two distinct points on the circle $|z| = r$. Let λ be a positive number. Let ϵ be an arbitrary positive number less than one. Then for $r = (1 - \epsilon)\pi/\lambda$*

$$\left| \frac{e^{-\lambda s_2} - e^{-\lambda s_1}}{\lambda s_2 - \lambda s_1} \right| > \frac{\epsilon}{(2 - \epsilon)^3} > \frac{\epsilon}{8} > 0. \tag{2.9}$$

To prove Lemma 3 we observe that if

$$F(\zeta) = \zeta + b_0 + \frac{b_1}{\zeta} + \dots + \frac{b_n}{\zeta^n} + \dots \tag{2.10}$$

is regular and schlicht for $|\zeta| > 1$, and if ζ_1 and ζ_2 are two distinct points for which $|\zeta_1| = |\zeta_2| = R > 1$ then it is known (**2**) that

$$\left|\frac{F(\zeta_2) - F(\zeta_1)}{\zeta_2 - \zeta_1}\right| > 1 - R^{-2}. \tag{2.11}$$

From (2.11) it follows that if

$$f(z) = z + b_2 z^2 + \ldots + b_n z^n + \ldots \tag{2.12}$$

is regular and univalent in $|z| < 1$, and if z_1 and z_2 are any two distinct points on $|z| = \rho < 1$, then

$$\left|\frac{f(z_2) - f(z_1)}{z_2 - z_1}\right| > \frac{1 - \rho}{(1 + \rho)^3}. \tag{2.13}$$

(2.13) follows from (2.11) if we define $F(\zeta) = \{f(\zeta^{-1})\}^{-1}$ and use the well-known inequality for univalent functions (2.12):

$$\left|\frac{f(z)}{z}\right| > (1 + \rho)^{-2},\ |z| = \rho < 1. \tag{2.14}$$

Since e^z is univalent in $|z| < \pi$, it follows from (2.13) that

$$\left|\frac{e^{z_2} - e^{z_1}}{z_2 - z_1}\right| > \frac{\pi^2(\pi - \rho)}{(\pi + \rho)^3},\ |z_1| = |z_2| = \rho < \pi \tag{2.15}$$

$$\left|\frac{e^{-\lambda s_2} - e^{-\lambda s_1}}{\lambda s_2 - \lambda s_1}\right| > \frac{\pi^2(\pi - \lambda r)}{(\pi + \lambda r)^3},\ |s_1| = |s_2| = r. \tag{2.16}$$

Choosing $r = (1 - \epsilon)\pi/\lambda$ in (2.16) we obtain (2.9). This completes the proof of Lemma 3.

3. Proof of Theorem 1. Let $s_0 = \sigma_0 + it_0$ be a complex number for which $\sigma_0 = \Re s_0 > \bar{\sigma} + (1 - \epsilon)\pi/\lambda_1$, $0 < \epsilon < 1$. Let s_1, s_2 be any two distinct values of s in the circle $|s - s_0| < r$ where $r < \sigma_0 - \bar{\sigma}$. Let $s_1' = s_1 - s_0$, $s_2' = s_2 - s_0$ so that $|s_i'| < r$. For an appropriate r we shall show that $f(s)$, given by the Dirichlet series (1.5) which is of class τ, is univalent in $|s - s_0| < r$, provided σ_0 is sufficiently large. In proving

$$\frac{f(s_2) - f(s_1)}{s_2 - s_1} \neq 0 \tag{3.1}$$

it will be sufficient to assume $|s_1'| = |s_2'| = r$. This follows from the fact that if the image curve of the circle $|s - s_0| = r$ by the mapping function bounds a simply connected region, the mapping function is schlicht in the interior when it is schlicht on the boundary. Choose $r = (1 - \epsilon)\pi/\lambda_1$. We now have

$$\begin{aligned} \frac{f(s_2) - f(s_1)}{s_2 - s_1} &= -\left(\frac{e^{-\lambda_1 s_2} - e^{-\lambda_1 s_1}}{s_2 - s_1}\right) + \sum_{n=2}^{\infty} a_n\left(\frac{e^{-\lambda_n s_2} - e^{-\lambda_n s_1}}{s_2 - s_1}\right) \\ &= -e^{-\lambda_1 s_0}\left(\frac{e^{-\lambda_1 s_2'} - e^{-\lambda_1 s_1'}}{s_2' - s_1'}\right) \\ &\qquad + \sum_{n=2}^{\infty} a_n e^{-\lambda_n s_0}\left(\frac{e^{-\lambda_n s_2'} - e^{-\lambda_n s_1'}}{s_2' - s_1'}\right) \end{aligned} \tag{3.2}$$

$$\left|\frac{f(s_2)-f(s_1)}{s_2-s_1}\right| \geqslant e^{-\lambda_1\sigma_0}\left|\frac{e^{-\lambda_1 s_2'}-e^{-\lambda_1 s_1'}}{s_2'-s_1'}\right| - R_q \tag{3.3}$$

where

$$R_q = \sum_{n=q}^{\infty} |a_n| e^{-\lambda_n\sigma_0}\left|\frac{e^{-\lambda_n s_2'}-e^{-\lambda_n s_1'}}{s_2'-s_1'}\right| \tag{3.4}$$

and a_q is the first non-vanishing coefficient a_n, $n \geqslant 2$. By Lemma 2, we have for $r < \sigma_0 - \tau \leqslant \sigma_0 - \bar{\sigma}$

$$\begin{aligned} R_q &\leqslant \sum_{n=q}^{\infty} \lambda_n |a_n| e^{\lambda_n(r-\sigma_0)} \\ &\leqslant \sum_{n=q}^{\infty} \lambda_n |a_n| e^{-\lambda_n\tau}\cdot e^{-\lambda_n(\sigma_0-r-\tau)} \\ &\leqslant e^{-\lambda_q(\sigma_0-r-\tau)}\cdot\sum_{n=2}^{\infty} \lambda_n|a_n|e^{-\lambda_n\tau} \\ &\leqslant e^{-\lambda_q(\sigma_0-r-\tau)}, \end{aligned} \tag{3.5}$$

where we have used the inequality (1.6) for functions $f(s)$ of class τ. From (3.3) and (3.5) we have for $r < \sigma_0 - \tau$.

$$\begin{aligned} \left|\frac{f(s_2)-f(s_1)}{s_2-s_1}\right| &\geqslant \lambda_1 e^{-\lambda_1\sigma_0}\left|\frac{e^{-\lambda_1 s_2'}-e^{-\lambda_1 s_1'}}{\lambda_1 s_2'-\lambda_1 s_1'}\right| - e^{-\lambda_q(\sigma_0-r-\tau)} \\ &\geqslant \lambda_1 e^{-\lambda_1\sigma_0}\cdot m(\lambda_1 r) - e^{-\lambda_q(\sigma_0-r-\tau)}, \end{aligned} \tag{3.6}$$

where

$$m(r) = \min_{|z_1|=|z_2|=r}\left|\frac{e^{z_2}-e^{z_1}}{z_2-z_1}\right| . \tag{3.7}$$

If $r < \pi$, $m(r) > 0$ since e^z is univalent in $|z| < \pi$. In spite of the fact that e^z is a simple elementary function, the problem of finding $m(r)$ as a function of r appears to be far from simple. It can be shown that

$$m(r) = \min_{0<x<r<\pi} e^{-(r^2-x^2)^{\frac{1}{2}}}\cdot\frac{\sin x}{x}. \tag{3.8}$$

We shall take $r = (1-\epsilon)\pi/\lambda_1$, where ϵ is an arbitrary number in the range $0 < \epsilon < 1$. We require the value of $m(\lambda_1, r) = m((1-\epsilon)\pi)$. For small values of ϵ,

$$m((1-\epsilon)\pi) \geqslant \frac{\epsilon}{\pi} + o(\epsilon).$$

However, we require a lower bound for $m((1-\epsilon)\pi)$ which holds uniformly for all ϵ in $0 < \epsilon < 1$. A positive lower bound of the correct order in ϵ is furnished in a simple way by the use of Lemma 3, which gives

$$m((1-\epsilon)\pi) \geqslant \frac{\epsilon}{(2-\epsilon)^3}, \qquad 0 < \epsilon < 1. \tag{3.9}$$

Thus, for $r = (1-\epsilon)\pi/\lambda_1$, $\sigma_0 > \tau + (1-\epsilon)\pi/\lambda_1$, we have

$$\left|\frac{f(s_2) - f(s_1)}{s_2 - s_1}\right| \geq e^{-\lambda_q\sigma_0}\{\lambda_1 e^{(\lambda_q-\lambda_1)\sigma_0} \,.\, m(\lambda_1 r) - e^{\lambda_q(\tau+r)}\}$$

$$(3.10) \qquad \geq e^{-\lambda_q\sigma_0}\left\{\frac{\lambda_1\epsilon}{(2-\epsilon)^3}\cdot e^{(\lambda_q-\lambda_1)\sigma_0} - e^{\lambda_q(\tau+(1-\epsilon)\pi/\lambda_1)}\right\}$$

$$> 0,$$

provided we choose σ_0 so that $\sigma_0 \geq \tau + (1 - \epsilon)\pi/\lambda_1$, and

$$(3.11) \qquad e^{(\lambda_q-\lambda_1)\sigma_0} > \frac{(2-\epsilon)^3}{\lambda_1\epsilon}\cdot e^{\lambda_q(\tau+(1-\epsilon)\pi/\lambda_1)},$$

that is

$$(3.12) \qquad \sigma_0 > \frac{3\log(2-\epsilon) - \log(\lambda_1\epsilon) + \lambda_q(\tau+(1-\epsilon)\pi/\lambda_1)}{\lambda_q - \lambda_1}.$$

We observe that the number π/λ_1, appearing in the radius $(1 - \epsilon)\pi/\lambda_1$ cannot be replaced by a larger one since the radius of univalency of the function

$$-e^{-\lambda_1 s},$$

which is the first term of the Dirichlet series (1.7), is exactly π/λ_1. We remark that for functions of the same class τ the value of α in (1.8) is independent of the function $f(s)$ once the sequence $\{\lambda_n\}$ has been selected. This completes the proof of Theorem 1.

It is by means of Theorem 1 that we are now able to establish the uniformly locally univalent property for all normalized Dirichlet series of the same class τ in a half-plane $\Re s > \beta$ where β has the value given in the following theorem.

THEOREM 2. *Let the class of functions* $\{f(s)\}$ *where*

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=q}^{\infty} a_n e^{-\lambda_n s}, \quad a_q \neq 0, s = \sigma + it,$$

be of the same class τ. *Then the functions* $f(s)$ *are uniformly locally univalent in the half-plane* $\Re s > \beta$ where

$$\beta = \begin{cases} \tau, & \text{if } \tau < \dfrac{\log\lambda_1}{\lambda_1}, \\ \dfrac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1} \geq \tau, & \text{if } \tau \geq \dfrac{\log\lambda_1}{\lambda_1}. \end{cases}$$

The functions $f(s)$ *of class* $(\log\lambda_1)/\lambda_1$ *are not uniformly locally univalent in* $\Re s > \tau - \eta$ *for arbitrarily small* $\eta > 0$, *and the functions* $f(s)$ *of class* $\tau > (\log\lambda_1)/\lambda_1$, *are not uniformly locally univalent in* $\Re s > \tau$ *while for an arbitrary* $\eta_1 > 0$, *the functions of a sub-class are uniformly locally univalent in* $\Re s > \tau + \eta_1$.

Before proving Theorem 2 we remark that Theorem 1 shows that the functions $f(s)$ of the same class τ are uniformly locally univalent in the half-plane $\Re s > \alpha - (1 - \epsilon)\pi/\lambda_1$ at least. As $\epsilon \to 1$ we have

$$\lim_{\epsilon\to 1}\alpha = \max\left(\tau, \frac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1}\right). \tag{3.13}$$

$$\lim_{\epsilon\to 1}\alpha = \tau \quad \text{if } \tau \leqslant \frac{\log\lambda_1}{\lambda_1}. \tag{3.14}$$

$$\lim_{\epsilon\to 1}\alpha = \frac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1} \quad \text{if } \tau \geqslant \frac{\log\lambda_1}{\lambda_1}. \tag{3.15}$$

In (1.8) we have $\alpha = \tau + (1 - \epsilon)\pi/\lambda_1$ provided

$$\tau \leqslant \frac{-1}{\lambda_1}\log\frac{(2-\epsilon)^3}{\lambda_1\epsilon} - \frac{(1-\epsilon)\pi}{\lambda_1}. \tag{3.16}$$

Suppose now that $\tau < (\log\lambda_1)/\lambda_1$. Then (3.16) is true for a range of ϵ,

$$0 < 1 - \frac{\delta\lambda_1}{\pi} \leqslant \epsilon < 1,$$

since

$$\frac{\log\lambda_1}{\lambda_1} \leqslant \frac{-1}{\lambda_1}\log\frac{(2-\epsilon)^3}{\lambda_1\epsilon} - \frac{(1-\epsilon)\pi}{\lambda_1} \tag{3.17}$$

for $\epsilon = 1$, but for no value of ϵ in the range $0 < \epsilon < 1$. For $\sigma_0 > \tau + \delta$, $f(s)$ is univalent in $|s - s_0| \leqslant \delta$ if $\tau < (\log\lambda_1)/\lambda_1$ since (3.16) is verified. Thus the functions $f(s)$ are uniformly locally univalent in $\Re s > \tau + \delta$ for arbitrarily small $\delta > 0$. It follows that the functions $f(s)$ are uniformly locally univalent in $\Re s > \tau$ whenever $\tau < (\log\lambda_1)/\lambda_1$.

Again, if $\tau \geqslant (\log\lambda_1)/\lambda_1$ we have

$$\alpha = [3\log(2-\epsilon) - \log(\lambda_1\epsilon) + \lambda_q\tau + (1-\epsilon)\pi\lambda_q/\lambda_1]/(\lambda_q - \lambda_1) \tag{3.18}$$

provided

$$\tau \geqslant -[3\log(2-\epsilon) - \log(\lambda_1\epsilon) + (1-\epsilon)\pi]/\lambda_1 \tag{3.19}$$

It is readily seen that (3.19) is verified for all ϵ, $0 < \epsilon < 1$, when $\tau \geqslant (\log\lambda_1)/\lambda_1$. Then for

$$\sigma_0 > \frac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1} + \delta, \qquad \delta > 0, \tag{3.20}$$

$f(s)$ is univalent in every circle $|s - s_0| \leqslant (1-\epsilon)\pi/\lambda_1$ provided

$$\frac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1} + \delta > \alpha \tag{3.21}$$

where α is given by (3.18). But since when $\epsilon = 1$, α has the value given by $\lim \alpha$ in (3.15) we see that for each given $\delta > 0$ there exists a range of values of ϵ, $0 < 1 - \delta_1 \leqslant \epsilon < 1$ for which (3.21) is verified. Since δ may be taken arbitrarily small it follows that the functions $f(s)$ are uniformly locally univalent for

$$\Re s > \frac{\lambda_q\tau - \log\lambda_1}{\lambda_q - \lambda_1} \text{ if } \tau \geqslant \frac{\log\lambda_1}{\lambda_1}. \tag{3.22}$$

If $\tau \geqslant (\log \lambda_1)/\lambda_1$, there exist functions $f(s)$ which are not locally univalent in $\Re s > \tau$ although they are locally univalent in $\Re s > \tau + \eta_1$ for a given $\eta_1 > 0$. For example, let $f(s)$, defined as

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=q}^{\infty} a_n e^{-\lambda_n s}, \qquad a_n > 0 \text{ for } n \geqslant q, \tag{3.23}$$

be of class $\tau \geqslant (\log \lambda_1)\lambda_1$ and choose q sufficiently large so that

$$\frac{\lambda_q \tau - \log \lambda_1}{\lambda_q - \lambda_1} < \tau + \eta_1, \qquad \eta_1 > 0. \tag{3.24}$$

The size of the coefficients a_n determine the value of τ,

$$f'(\tau) = \lambda_1 e^{-\lambda_1 \tau} - \sum_{n=q}^{\infty} \lambda_n a_n e^{-\lambda_n \tau} = \lambda_1 e^{-\lambda_1 \tau} - 1. \tag{3.25}$$

If $\tau = (\log \lambda_1)/\lambda_1, f'(\tau) = 0$. If $\tau > (\log \lambda_1)/\lambda_1, f'(\tau) < 0$, whereas $f'(\sigma) > 0$ for large values of σ, since

$$\lim_{\sigma \to +\infty} e^{\lambda_1 \sigma} f'(\sigma) = \lambda_1 > 0.$$

Thus $f'(\sigma)$, being continuous, must vanish for at least one value $\sigma = \sigma_1 > \tau$. But $\sigma_1 < \tau + \eta_1$ since $f(s)$ is locally univalent at least for $\Re s > \tau + \eta_1$, and $f'(\sigma)$ can not vanish for $\sigma > \tau + \eta_1$. It follows that $f(s)$ is not schlicht in the neighbourhood of σ_1. Thus, if τ exceeds $(\log \lambda_1)/\lambda_1$, $f(s)$ need not be locally univalent in $\Re s > \tau$ even though it is for $\Re s > \tau + \eta_1$. It is also seen that if $\tau = (\log \lambda_1)/\lambda_1$, $f(s)$ need not be locally univalent in $\Re s > \tau - \eta$, $\eta > 0$. This completes the proof of Theorem 2.

We shall now make an application of Theorem 2 to the Riemann zeta-function $\zeta(s)$.

$$1 - \zeta(s) = -\sum_{n=1}^{\infty} e^{-s \log(n+2)} = -\sum_{n=1}^{\infty} (n+1)^{-s}. \tag{3.26}$$

Now $1 - \zeta(s)$ is of class τ where

$$\sum_{n=2}^{\infty} \frac{\log(n+1)}{(n+1)^{\tau}} = 1, \quad \zeta'(\tau) + 2^{-\tau} \log 2 + 1 = 0. \tag{3.27}$$

Since

$$\frac{\log \lambda_1}{\lambda_1} = \frac{\log \log 2}{\log 2} < 0$$

and $\tau > 1$, we see that $\zeta(s)$ is locally univalent for

$$\Re s > \frac{\tau \lambda_2 - \log \lambda_1}{\lambda_2 - \lambda_1} = \frac{\tau \log 3 - \log \log 2}{\log 3 - \log 2} = 2.70749\,\tau + 0.90428, \tag{3.28}$$

where τ is the solution of the equation (3.27). Since

$$\sum_{n=2}^{\infty} \frac{\log(n+1)}{(n+1)^{\tau}} < \int_2^{\infty} \frac{\log x}{x^{\tau}}\, dx = \frac{(\tau - 1)\log 2 + 1}{(\tau - 1)^2\, 2^{\tau - 1}} = 1 \tag{3.29}$$

for a value $\tau = \tau_0$ in the range $1.9 < \tau_0 < 2.0$ it follows that $1 - \zeta(s)$ is of class $\tau < 2$. Also, since

$$\sum_{n=2}^{\infty} \frac{\log (n+1)}{(n+1)^\tau} > \frac{\log 3}{3^\tau} + \int_3^\infty \frac{\log x}{x^\tau}\, dx = \frac{\log 3}{3^\tau} + \frac{(\tau-1)\log 4 + 1}{(\tau-1)^2 4^{\tau-1}} = 1 \tag{3.30}$$

for a value of $\tau = \tau_1$ in the range $1.9 < \tau_1 < 2.0$ it follows that $1 - \zeta(s)$ is of class $\tau > 1.9$. Hence, the class of $1 - \zeta(s)$ lies in the range $1.9 < \tau < 2.0$. We conclude that $\zeta(s)$ is locally schlicht in a half-plane $\Re s > c$ where $c < 6.32$.

4. Proof of Theorem 3. Univalency in strips. Instead of examining $f(s)$, given by (1.5) and of class τ, for univalency in circles $|s - s_0| < r$, we shall turn now to a similar task for strips. Let D_k denote the strip of the $s = \sigma + it$ plane defined by $\sigma > \tau$, where $\tau < \tau_0$ in the notation of (1.10), and $-t_0 < t - 2k\pi/\lambda_1 < t_0$, where k is an arbitrary integer and

$$t_0 = \frac{1}{\lambda_1} \operatorname{arc\,cos}\left(\frac{e^{\lambda_1 \tau}}{\lambda_1}\right), \qquad 0 < t_0 < \pi/2\lambda_1. \tag{4.1}$$

Let C_k denote the boundary of D_k and consist of the three line segments α_k, β_k, γ_k defined as follows.

α_k: that part of C_k which lies on $t = t_0 + 2k\pi/\lambda_1$,

β_k: that part of C_k which lies on $\sigma = \tau$,

γ_k: that part of C_k which lies on $t = -t_0 + 2k\pi/\lambda_1$.

Let D^*_k denote the rectangular sub-domain of D_k whose boundary C^*_k consists of the parts of the two line segments α_k and γ_k for which $\tau < \sigma < \tau^*$, together with β_k and δ_k, where δ_k denotes the line segment $\sigma = \tau^* > \tau$, $-t_0 + 2k\pi/\lambda_1 < t < t_0 + 2k\pi/\lambda_1$.

We shall show that $f(s)$ is univalent in the domains D_k and that $w = f(s)$ maps C_k onto a simple, closed Jordan curve Γ_k which is convex in the direction of the real axis, which is to say that the region bounded by Γ_k is star-shaped with respect to the point at infinity at an end of the real axis. Since $\lim_{\sigma \to +\infty} f(\sigma) = 0$, it follows that the only zero $f(s)$ has in D_k corresponds to the point of D_k at infinity. Thus Γ_k passes through the origin in the w-plane. If w_1 and w_2 are any two distinct points of the image of D_k by $w = f(s)$ for which $\Im w_1 = \Im w_2$, it will follow that the line segment joining w_2 and w_1 lies entirely within the region encompassed by Γ_k. If w_1 and w_2 are any two points interior to Γ_k, they must also lie interior to Γ^*_k, the image of C^*_k, if τ^* is taken sufficiently large. Thus it is sufficient to prove that the region bounded by Γ^*_k is convex in the direction of the real axis for every $\tau^* > \tau$.

On β_k we have $\sigma = \tau$ and

$$f(\tau + it) = -e^{-\lambda_1(\tau+it)} + \sum_{n=2}^{\infty} a_n e^{-\lambda_n(\tau+it)}. \tag{4.2}$$

Because $f(s)$ is of class τ, $f(s)$ and $f'(s)$ are continuous on $\sigma = \tau \geqq \bar{\sigma}$, and we have

$$\mathfrak{I}f(\tau + it) = e^{-\lambda_1\tau} \sin(\lambda_1 t) - \sum_{n=2}^{\infty} \{\alpha_n \sin(\lambda_n t) - \beta_n \cos(\lambda_n t)\} e^{-\lambda_n\tau} \tag{4.3}$$

where $a_n = \alpha_n + i\beta_n$, α_n and β_n real, and

$$\begin{aligned}\frac{\partial}{\partial t}\mathfrak{I}f(\tau + it) &= \lambda_1 e^{-\lambda_1\tau} \cos(\lambda_1 t) \\ &\quad - \sum_{n=2}^{\infty} \lambda_n\{\alpha_n \cos(\lambda_n t) + \beta_n \sin(\lambda_n t)\} e^{-\lambda_n\tau}.\end{aligned} \tag{4.4}$$

Since

$$|\alpha_n \cos\theta + \beta_n \sin\theta| \leqq (\alpha_n^2 + \beta_n^2)^{\frac{1}{2}} = |a_n| \tag{4.5}$$

for all θ we have

$$\begin{aligned}\frac{\partial}{\partial t}\mathfrak{I}f(\tau + it) &\geqq \lambda_1 e^{-\lambda_1 t} \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n\tau} \\ &\geqq \lambda_1 e^{-\lambda_1 t} \cos(\lambda_1 t) - 1 \geqq 0\end{aligned} \tag{4.6}$$

for

$$-t_0 \leqq t - \frac{2k\pi}{\lambda_1} \leqq t_0,\ 0 < t_0 = \frac{1}{\lambda_1} \operatorname{arc} \cos\left(\frac{e^{\lambda_1\tau}}{\lambda_1}\right) \leqq \frac{\pi}{2\lambda_1}.$$

Thus, $\mathfrak{I}f(s)$ is a monotonically increasing function of t on β_k.

A similar proof holds on δ_k where $\sigma = \tau^* > \tau$ with a slight modification. Here we have

$$\begin{aligned}\frac{\partial}{\partial t}\mathfrak{I}f(\tau^* + it) &\geqq \lambda_1 e^{-\lambda_1\tau^*} \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n\tau^*} \\ &= e^{-\lambda_1\tau^*}\left\{\lambda_1 \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-(\lambda_n-\lambda_1)\tau^*}\right\} \\ &\geqq e^{-\lambda_1\tau^*}\left\{\lambda_1 \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-(\lambda_n-\lambda_1)\tau}\right\} \\ &\geqq e^{-\lambda_1\tau^*}\{\lambda_1 \cos(\lambda_1 t) - e^{\lambda_1\tau}\} \geqq 0\end{aligned} \tag{4.7}$$

for $|t - 2k\pi/\lambda_1| \leqq t_0$. Thus $\mathfrak{I}f(s)$ is a monotonically increasing function of t on δ_k.

On α_k we have $t = t_k = t_0 + 2k\pi/\lambda_1$, $\tau < \sigma < \tau^*$, and

$$\mathfrak{J}f(\sigma + it_k) = e^{-\lambda_1\sigma}\sin(\lambda_1 t_0) - \sum_{n=2}^{\infty}\{\alpha_n\sin(\lambda_n t_k) - \beta_n\cos(\lambda_n t_k)\}e^{-\lambda_n\sigma}, \tag{4.8}$$

$$\begin{aligned}
\frac{\partial\mathfrak{J}}{\partial\sigma}f(\sigma + it_k) &= -\lambda_1 e^{-\lambda_1\sigma}\sin(\lambda_1 t_0) \\
&\qquad + \sum_{n=2}^{\infty}\lambda_n\{\alpha_n\sin(\lambda_n t_k) - \beta_n\cos(\lambda_n t_k)\}e^{-\lambda_n\sigma} \\
&< -\lambda_1 e^{-\lambda_1\sigma}\sin(\lambda_1 t_0) + \sum_{n=2}^{\infty}\lambda_n|a_n|e^{-\lambda_n\sigma} \\
&< e^{-\lambda_1\sigma}\left\{-\lambda_1\sin(\lambda_1 t_0) + \sum_{n=2}^{\infty}\lambda_n|a_n|e^{-(\lambda_n-\lambda_1)\sigma}\right\} \\
&< e^{-\lambda_1\sigma}\left\{-\lambda_1\sin(\lambda_1 t_0) + \sum_{n=2}^{\infty}\lambda_n|a_n|e^{-(\lambda_n-\lambda_1)\tau}\right\} \\
&< e^{-\lambda_1\sigma}\{-\lambda_1\sin(\lambda_1 t_0) + e^{\lambda_1\tau}\} \\
&= e^{-\lambda_1\sigma}\{-(\lambda_1^2 - e^{2\lambda_1\tau})^{\frac{1}{2}} + e^{\lambda_1\tau}\} < 0,
\end{aligned} \tag{4.9}$$

provided, in the notation of (1.10),

$$\tau < \tau_0, \qquad \sigma > \tau. \tag{4.10}$$

Thus, $\mathfrak{J}f(s)$ is a monotonically decreasing function of σ on α_k.

A similar argument shows that $\mathfrak{J}f(s)$ is a monotonically increasing function of σ on γ_k (t_0 is replaced by $-t_0$).

Combining the above results we have shown that, as three sides of the rectangle C^*_k are traversed in the counter-clockwise direction beginning at the point of intersection of β_k and γ_k and ending at the point of intersection of α_k and β_k, the corresponding arc of the curve Γ^*_k has the property that every horizontal straight line (parallel to the real axis) cuts it in at most one point since $\mathfrak{J}f(s)$ is non-decreasing. Similarly, the image of β_k also has the property that every horizontal line cuts it in at most one point. Thus, the region bounded by Γ^*_k is convex in the direction of the real axis for every $\tau^* > \tau$. Since Γ^*_k has therefore no double points $f(s)$ must be univalent in D^*_k, and consequently univalent in D_k as well.

We next see that there exist functions $f(s)$ and certain sequences $\{\lambda_n\}$ for which the theorem is not true if the strip D_k is enlarged by keeping the sides parallel to the axes of reference. Let $\epsilon > 0$ be chosen arbitrarily. Choose λ_2 so that

$$e^{(\lambda_2-\lambda_1)\epsilon} > 2^{\frac{1}{2}}, \quad \lambda_2 > \lambda_1.$$

Choose the coefficients a_n of (1.9) positive and so that $f(s)$ is of class $\tau = \tau_0$. Then for $\sigma = \tau - \epsilon$, $t = 0$ we have

$$\frac{\partial}{\partial t}\Im f(\tau - \epsilon + it) = \lambda_1 e^{-\lambda_1(\tau-\epsilon)} - \sum_{n=2}^{\infty} \lambda_n a_n e^{-\lambda_n(\tau-\epsilon)} \tag{4.11}$$

$$< \lambda_1 e^{-\lambda_1\tau}. e^{\lambda_1\epsilon} - e^{\lambda_2\epsilon}$$

$$= 2^{\frac{1}{2}} e^{\lambda_1\epsilon} - e^{\lambda_2\epsilon} < 0.$$

Thus $\Im f(s)$ in this case is not steadily increasing as t increases on $\tau = \tau_0 - \epsilon$. This shows that we cannot enlarge the strip D_0 horizontally and have Theorem 3 valid for all functions $f(s)$ of the class considered.

Next we shall show that the strip may not be enlarged vertically. Choose $\lambda_1 > 0$ and $\epsilon > 0$ arbitrarily, and, for $n \geqslant 2$, choose $\lambda_n = (4n+1)\pi/(2t_0 + 2\epsilon)$ where t_0 is defined as in (1.11) and where

$$\tau < \tau_0.$$

We choose the coefficients a_n of (1.9) so that for $n \geqslant 2$, $\alpha_n = \Re a_n = 0$, $\beta_n = \Im a_n > 0$, with a proper choice of magnitude of β_n so that $f(s)$ is of the given class τ. Then for $t = t_0 + \epsilon$, $\epsilon > 0$, ϵ small, and $\sigma = \tau$,

$$\frac{\partial}{\partial t}\Im f(\tau + it)\Big]_{t=t_0+\epsilon} = \lambda_1 e^{-\lambda_1\tau} \cos \lambda_1(t_0 + \epsilon) - \sum_{n=2}^{\infty} \lambda_n \beta_n e^{-\lambda_n\tau} \tag{4.12}$$

$$\leqslant \lambda_1 e^{-\lambda_1\tau} \cos \lambda_1(t_0 + \epsilon) - 1$$

$$< \lambda_1 e^{-\lambda_1\tau} \cos (\lambda_1 t_0) - 1$$

$$= 0.$$

It follows that $\Im f(\tau + it)$ is not monotonically increasing for $|t| < t_0 + \epsilon$, $\epsilon > 0$, for this function of class τ. Therefore the strip D_0 cannot be enlarged vertically for all functions considered in Theorem 3.

It is also seen that no larger value of τ than the one given by (1.10) is permissible. Theorem 3 is, therefore, a "best possible" one. This completes the proof of the first part of Theorem 3 where τ is restricted as in (1.10).

On the other hand, we may increase the range of τ slightly if mere univalency is demanded in the strips D_k. We assume that the inequality $\tau \leqslant \tau_0$ where τ_0 is defined in (1.10) is replaced by $\tau \leqslant \tau^*_0 = (\log \lambda_1)/\lambda_1$ and shall show that $f(s)$ is still univalent in D_k. We make use of Noshiro's Theorem **(4)** and show first that $\Re f'(s) > 0$ in D_k. Since

$$\Re f'(s) = \lambda_1 e^{-\lambda_1\sigma} \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n \{\alpha_n \cos(\lambda_n t) - \beta_n \sin(\lambda_n t)\} e^{-\lambda_n\sigma}, \tag{4.13}$$

$$\Re f'(s) \geqslant \lambda_1 e^{-\lambda_1\sigma} \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n\sigma}$$

$$\geqslant e^{-\lambda_1\sigma}\left\{\lambda_1 \cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-(\lambda_n-\lambda_1)\sigma}\right\} \tag{4.14}$$

$$\geqslant e^{-\lambda_1\sigma}\{\lambda_1 \cos(\lambda_1 t) - e^{\lambda_1\tau}\} > 0$$

if s is in D_k and $\tau < \tau^*_0$.

Since we have shown that $\Re f'(s) > 0$ in D_k and since D_k is convex it follows at once by Noshiro's Theorem that $f(s)$ is univalent in D_k.

No larger value than τ^*_0 for τ is permissible in general. Indeed if $a_n > 0$ for $n \geqslant 2$, and $\tau > \tau^*_0$ we have

$$f'(\tau) = e^{-\lambda_1\tau}(\lambda_1 - e^{\lambda_1\tau}) < 0 \tag{4.15}$$

if

$$\sum_{n=2}^{\infty} \lambda_n a_n e^{-\lambda_n\tau} = 1. \tag{4.16}$$

But

$$\begin{aligned} f'(\sigma) &= \lambda_1 e^{-\lambda_1\sigma} - \sum_{n=2}^{\infty} \lambda_n a_n e^{-\lambda_n\sigma} \\ &\geqslant 0 \end{aligned} \tag{4.17}$$

for σ sufficiently large. Thus $f'(s)$ vanishes in the strip D_0 in this case. In this case $f(s)$ is not univalent.

It should be noticed also that if $\tau < \tau^*_0$ then

$$\frac{\partial}{\partial\sigma}\Re f(\sigma + it) \geqslant 0$$

on t = constant, $|t| \leqslant t_0$, $\sigma \geqslant \tau$. For

$$\begin{aligned} \frac{\partial}{\partial\sigma}\Re f(\sigma + it) &= \lambda_1 e^{-\lambda_1\sigma}\cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n\{\alpha_n \cos(\lambda_n t) - \beta_n \sin(\lambda_n t)\} e^{-\lambda_n\sigma} \\ &\geqslant \lambda_1 e^{-\lambda_1\sigma}\cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-\lambda_n\sigma} \\ &= e^{-\lambda_1\sigma}\left\{\lambda_1\cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-(\lambda_n-\lambda_1)\sigma}\right\} \\ &\geqslant e^{-\lambda_1\sigma}\left\{\lambda_1\cos(\lambda_1 t) - \sum_{n=2}^{\infty} \lambda_n |a_n| e^{-(\lambda_n-\lambda_1)\tau}\right\} \\ &\geqslant e^{-\lambda_1\sigma}\{\lambda_1\cos(\lambda_1 t) - e^{\lambda_1\tau}\} \\ &\geqslant 0 \end{aligned} \tag{4.18}$$

for $|t| \leqslant t_0$, $\tau < \tau^*_0$. Thus, if $\tau < \tau^*_0$, D_k is mapped into Δ_k by $w = f(s)$, and part of the boundary of Δ_k is convex in the direction of the imaginary axis while the remaining part of the boundary is convex in the direction of the real axis. These parts correspond to sides of D_k regardless of which function $f(s)$ of class τ is used.

We have completed the proof of Theorem 3, and the following corollary is a consequence of the preceding remarks.

COROLLARY 1. *If $\{f_n(s)\}$ is a sequence of functions defined by Dirichlet series*

$$f_n(s) = -e^{-\lambda_1 s} + \sum_{m=2}^{\infty} a_m^{(n)} e^{-\lambda_m^{(n)} s}, \tag{4.19}$$

relative to the sequence $\{\lambda_m^{(n)}\}$,

$$0 < \lambda_1 < \lambda_2^{(n)} < \lambda_3^{(n)} < \ldots < \lambda_m^{(n)} < \ldots, \qquad \lambda_m^{(n)} \to \infty, \tag{4.20}$$

and if each $f_n(s)$ *is of the same class* τ, *so that*

$$\sum_{m=2}^{\infty} \lambda_m^{(n)} |a_m^{(n)}| e^{-\lambda_m^{(n)} \tau} < 1, \quad \tau < \frac{\log \lambda_1}{\lambda_1}, \tag{4.21}$$

then, for each sequence $\{A_n\}$ *of positive real numbers for which*

$$\phi(s) = \sum_{n=1}^{\infty} A_n f_n(s) \tag{4.22}$$

converges uniformly to $\phi(s)$ *in* $\Re s > \alpha$, $\alpha < \tau$, *we have* $\phi(s)$ *analytic in* $\Re s > \alpha$, *and* $\phi(s)$ *is univalent in each strip* D_k *of Theorem* 3.

COROLLARY 2. *Let*

$$f(s) = -e^{-\lambda_1 s} + \sum_{n=2}^{\infty} a_n e^{-\lambda_n s}, \qquad s = \sigma + it, \tag{4.23}$$

be absolutely convergent for $\sigma > \bar{\sigma}$, $-\infty < \bar{\sigma} < \infty$, *and let* $f(s)$ *be of class* $\tau < (\log \lambda_1)/\lambda_1$. *Let*

$$t_0 = \frac{1}{\lambda_1} \operatorname{arc\,cos}\left(\frac{e^{\lambda_1 \tau}}{\lambda_1}\right), \quad 0 < t_0 < \frac{\pi}{2\lambda_1}. \tag{4.24}$$

Then $f(s)$ *is univalent in every semi-infinite strip* D *of width* $2t_0$ *which is parallel to the real axis and lies in the half-plane* $\Re s \geqslant \tau$.

In order to see that Corollary 2 follows from Theorem 3 we observe that if t' is an arbitrary real number the function

$$F(s) = e^{i\lambda_1 t'} . f(s + it') = -e^{-\lambda_1 s} + \sum_{n=2}^{\infty} a_n e^{-i(\lambda_n - \lambda_1) t'} \cdot e^{-\lambda_n s} \tag{4.25}$$

is a Dirichlet series of the same class τ as the class of $f(s)$. Applying Theorem 3 to $F(s)$ we find that $F(s)$ is univalent in each strip D_k of width $2t_0$. Hence $f(s)$ is univalent in a strip obtained by a translation vertically of the strip D_k by the arbitrary value t'.

If $\tau < \tau_0$ as in (1.10), we can conclude further that $f(s)$ is convex in some one direction in each strip of width $2t_0$, $\sigma > \tau$, parallel to the real axis. The direction of convexity varies with the position of each strip in general. As in Theorem 3 Corollary 2 is sharp.

References

1. J. W. Alexander, *Functions which map the interior of the unit circle upon simple regions*, Ann. of Math., *17* (1915), 12–22.
2. G. M. Goluzin, *On distortion theorems and coefficients of univalent functions*, Rec. Math. (Mat. Sbornik), N.S. (61), *19* (1946), 183–202.
3. P. Montel, *Sur les fonctions localement univalentes ou multivalentes*, Ann. Sci. École Norm. Sup. (3), *54* (1937), 39–54.
4. J. Noshiro, *On the theory of schlicht functions*, J. Fac. Sci. Hokkaido Univ. (1), *2* (1934), 129–155.
5. R. Remak, *Ueber eine specielle Klasse schlichter konformer Abbildungen des Einheitskreises*, Mathematica, Zutphen, *11* (1943), 175–192; *12* (1943), 43–49.

Rutgers, the State University of New Jersey.

LAPLACE TRANSFORMS AND GENERALIZED LAGUERRE POLYNOMIALS

P. G. ROONEY

1. Introduction. Various sets of necessary and sufficient conditions are known in order that a function $f(s)$, analytic for Re $s > 0$, be represented as the Laplace transform of a function in $L_p(0, \infty)$, $1 < p < \infty$. Most of these theories are based on the properties of some inversion operator for the transformation—see, for example, (**7**, chap. 7). However in the case $p = 2$ a number of representation theorems of a much simpler type are available. One of these is due to Shohat (**5**) who has in effect shown that a necessary and sufficient condition for such a representation, with $p = 2$, is that

$$\sum_{n=0}^{\infty} |q_n|^2 < \infty,$$

where

$$q_n = \sum_{r=0}^{n} \binom{n}{r} \frac{1}{r!} f^{(r)}(\tfrac{1}{2}).$$

Shohat's proof makes use of the Laguerre polynomials.

Recently the author has given (**4**) necessary and sufficient conditions that $f(s)$ be the Laplace transform of a function of the form $t^{\lambda}F(t)$, $F \in L_p(0, \infty)$, $1 < p < \infty$, $\lambda > -1/q$, where $p^{-1} + q^{-1} = 1$. These conditions were given in terms of a particular inversion operator. In this paper we shall see that Shohat's theorem can be generalized, for $p = 2$, to cover this more general case. This is done in § 2 below, using generalized Laguerre polynomials. We also obtain there an expression for $F(t)$ which we shall use in § 3 to obtain some results about Hankel transforms. For convenience we write $\lambda = \frac{1}{2}\nu$ throughout the following.

2. Representation theorem. We start with a preliminary lemma.

LEMMA 1. *If $f(s)$ is analytic for* Re $s > 0$ *and $\nu > -1$, then*

$$f(s) = \frac{1}{(s + \frac{1}{2})^{\nu+1}} \sum_{n=0}^{\infty} q_n \left(\frac{s - \frac{1}{2}}{s + \frac{1}{2}}\right)^n,$$

where

$$q_n = \sum_{r=0}^{n} \binom{n + \nu}{n - r} \frac{1}{r!} f^{(r)}(\tfrac{1}{2}),$$

the branch of $(s + \frac{1}{2})^{\nu+1}$ that is positive when $s + \frac{1}{2}$ is positive being chosen.

Received May 31, 1957. This work was done in part while the author was the holder of a summer associateship of the National Research Council of Canada.

Proof. Let

$$s = \tfrac{1}{2}\frac{1+z}{1-z},$$

and $f(s) = F(z)$. Then $F(z)$ is analytic in $|z| < 1$, and hence so is $F(z)/(1-z)^{\nu+1}$. Thus

$$F(z)/(1-z)^{\nu+1} = \sum_{n=0}^{\infty} q_n z^n, \qquad |z| < 1,$$

where if $r < 1$

$$\begin{aligned} q_n &= \frac{1}{2\pi i}\int_{|z|=r} (F(z)/z^{n+1}\,(1-z)^{\nu+1})dz \\ &= \text{Residue}_{z=0}\,(F(z)/z^{n+1}(1-z)^{\nu+1}) \\ &= \text{Residue}_{s=\frac{1}{2}}\,(f(s)\,(s+\tfrac{1}{2})^{n+\nu}/(s-\tfrac{1}{2})^{n+1}) \\ &= \frac{1}{n!}\lim_{s\to\frac{1}{2}}\left\{\frac{d^n}{ds^n}\,(f(s)\,(s+\tfrac{1}{2})^{n+\nu})\right\} \\ &= \frac{1}{n!}\lim_{s\to\frac{1}{2}}\left\{\sum_{r=0}^{n}\binom{n}{r} f^{(r)}(s)\,\frac{\Gamma(n+\nu+1)}{\Gamma(r+\nu+1)}\,(s+\tfrac{1}{2})^{r+\nu}\right\} \\ &= \sum_{r=0}^{n}\binom{n+\nu}{n-r}\frac{1}{r!}f^{(r)}(\tfrac{1}{2}). \end{aligned}$$

Hence

$$f(s) = F(z) = (1-z)^{\nu+1}\sum_{n=0}^{\infty} q_n z^n = \frac{1}{(s+\frac{1}{2})^{\nu+1}}\sum_{n=0}^{\infty} q_n\left(\frac{s-\frac{1}{2}}{s+\frac{1}{2}}\right)^n.$$

THEOREM 1. *A necessary and sufficient condition that a function $f(s)$, analytic for* Re $s > 0$, *be the Laplace transform of a function of the form $t^{\frac{1}{2}\nu}F(t)$, with $F \in L_2(0, \infty)$ and $\nu > -1$, is that*

$$\sum_{n=0}^{\infty}\frac{n!}{\Gamma(\nu+n+1)}\,|q_n|^2 < \infty$$

where

$$q_n = \sum_{r=0}^{n}\binom{n+\nu}{n-r}\frac{1}{r!}f^{(r)}(\tfrac{1}{2}).$$

In this case

$$F(t) = \operatorname*{l.i.m.}_{r\to\infty} t^{\frac{1}{2}\nu}\,e^{-\frac{1}{2}t}\sum_{n=0}^{r}\frac{n!}{\Gamma(\nu+n+1)}\,q_n\,L_n^{(\nu)}(t),$$

and

$$\sum_{n=0}^{\infty}\frac{n!}{\Gamma(\nu+n+1)}\,|q_n|^2 = \int_0^{\infty}|F(t)|^2\,dt.$$

Proof of necessity. Suppose

$$f(s) = \int_0^\infty e^{-st} t^{\frac{1}{2}\nu} F(t)dt, \qquad F \in L_2(0, \infty), \nu > -1.$$

Let

$$\phi_n(t) = \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} e^{-\frac{1}{2}t} t^{\frac{1}{2}\nu} L_n^{(\nu)}(t).$$

Then, as is well known, $\{\phi_n\}$ is a complete orthonormal sequence in $L_2(0, \infty)$. We have, using **(2**, §10.12(7)**)** and **(1**, chap. 3, §2**)**,

$$\begin{aligned}(F, \phi_n) &= \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} \int_0^\infty e^{-\frac{1}{2}t} t^{\frac{1}{2}\nu} L_n^{(\nu)}(t)\, F(t)\, dt \\ &= \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} \sum_{r=0}^{n} \binom{n+\nu}{n-r} \frac{1}{r!} \int_0^\infty e^{-\frac{1}{2}t}(-t)^r t^{\frac{1}{2}\nu} F(t)dt \\ &= \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} \sum_{r=0}^{\infty} \binom{n+\nu}{n-r} \frac{1}{r!} f^{(r)}(\tfrac{1}{2}) = \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} q_n.\end{aligned}$$

Hence

$$\begin{aligned}F(t) &= \operatorname*{l.i.m.}_{r\to\infty} \sum_{n=0}^{r} (F, \phi_n)\, \phi_n(t) \\ &= \operatorname*{l.i.m.}_{r\to\infty} t^{\frac{1}{2}\nu} e^{-\frac{1}{2}t} \sum_{n=0}^{r} \frac{n!}{\Gamma(\nu + n + 1)} q_n L_n^{(\nu)}(t),\end{aligned}$$

and from the Parseval relation

$$\sum_{n=0}^\infty \frac{n!}{\Gamma(\nu + n + 1)} |q_n|^2 = \sum_{n=0}^\infty |(F, \phi_n)|^2 = \int_0^\infty |F(t)|^2\, dt < \infty .$$

Proof of sufficiency. Since

$$\sum_{n=0}^\infty \frac{n!}{\Gamma(\nu + n + 1)} |q_n|^2 < \infty ,$$

by the Riesz-Fischer theorem there is a function $F \in L_2(0, \infty)$ such that

$$(F, \phi_n) = q_n \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}}.$$

Let $G(t) = t^{\frac{1}{2}\nu} e^{-\bar{s}t}$, Re $s > 0$. Then $G \in L_2(0, \infty)$, and from **(3**, §4.11(28)**)**,

$$\begin{aligned}(G, \phi_n) &= \int_0^\infty t^{\frac{1}{2}\nu} e^{-\bar{s}t} \phi_n(t)\, dt \\ &= \left(\frac{n!}{\Gamma(\nu + n + 1)}\right)^{\frac{1}{2}} \int_0^\infty e^{-(\bar{s}+\frac{1}{2})t} t^\nu L_n^{(\nu)}(t)\, dt \\ &= \left(\frac{\Gamma(\nu + n + 1)}{n!}\right)^{\frac{1}{2}} \frac{(\bar{s} - \frac{1}{2})^n}{(\bar{s} + \frac{1}{2})^{n+\nu+1}}.\end{aligned}$$

Hence from Lemma 1 and Parseval's relation, if Re $s > 0$,

$$\begin{aligned} f(s) &= \frac{1}{(s+\frac{1}{2})^{\nu+1}} \sum_{r=0}^{\infty} q_n \left(\frac{s-\frac{1}{2}}{s+\frac{1}{2}}\right)^n \\ &= \sum_{n=0}^{\infty} \left\{ q_n \left(\frac{n!}{\Gamma(\nu+n+1)}\right)^{\frac{1}{2}} \right\} \left\{ \left(\frac{\Gamma(\nu+n+1)}{n!}\right)^{\frac{1}{2}} \frac{(s-\frac{1}{2})^n}{(s+\frac{1}{2})^{n+\nu+1}} \right\} \\ &= \sum_{n=0}^{\infty} (F, \phi_n)\, \overline{(G, \phi_n)} = (F, G) = \int_0^{\infty} e^{-st}\, t^{\frac{1}{2}\nu} F(t)\, dt. \end{aligned}$$

3. Application to Hankel transforms. For our purposes here we shall define the Hankel transform for $F \in L_2(0, \infty)$, $\nu > -1$, by

$$G(x) = \frac{d}{dx} \int_0^{\infty} k_\nu(xy)\, F(y) \frac{dy}{y}$$

where

$$k_\nu(x) = \int_0^x J_\nu(2\sqrt{y})\, dy.$$

Since the Mellin transform of $J_\nu(2\sqrt{y})$ is

$$\Gamma(s + \tfrac{1}{2}\nu)/\Gamma(\tfrac{1}{2}\nu - s + 1), \qquad -\tfrac{1}{2}\nu < \text{Re } s < 3/4,$$

it follows since $\nu > -1$ that the hypotheses of (**6**, Theorem 129) are satisfied so that if $F \in L_2(0, \infty)$, G exists and is in $L_2(0, \infty)$, and Parseval's equation holds. Further

$$F(x) = \frac{d}{dx} \int_0^{\infty} k_\nu(xy)\, G(y) \frac{dy}{y}.$$

Here we shall use the results of Theorem 1 to invert the Hankel transform. We first prove the following lemma (compare (**1**, chap 2 §16)).

LEMMA 2. *If* $F \in L_2(0, \infty)$, $\nu > -1$,

$$G(x) = \frac{d}{dx} \int_0^{\infty} k_\nu(xy)\, F(y)\, dy$$

where

$$k_\nu(x) = \int_0^x J_\nu(2\sqrt{y})\, dy,$$

$$f(s) = \int_0^{\infty} e^{-st}\, t^{\frac{1}{2}\nu} F(t) dt$$

and

$$g(s) = \int_0^{\infty} e^{-st}\, t^{\frac{1}{2}\nu} G(t) dt,$$

then

$$f(s) = \frac{1}{s^{\nu+1}} g(1/s).$$

Proof. The Hankel transform of $t^{\frac{1}{2}\nu}e^{-st}$ is given, on using (3, §4.14(30)), by

$$\begin{aligned}\frac{d}{dx}\int_0^\infty t^{\frac{1}{2}\nu-1}\, e^{-st}\, dt \int_0^{xt} J_\nu(2\sqrt{y})dy\\ &= \frac{d}{dx}\int_0^\infty t^{\frac{1}{2}\nu}\, e^{-st}\, dt\int_0^x J_\nu(2\sqrt{yt})dy\\ &= \frac{d}{dx}\int_0^x dy\int_0^\infty t^{\frac{1}{2}\nu}\, e^{-st}\, J_\nu(2\sqrt{yt})dt\\ &= \int_0^\infty t^{\frac{1}{2}\nu}\, e^{-st}\, J_\nu(2\sqrt{xt})dt\\ &= \frac{x^{\frac{1}{2}\nu}\, e^{-x/s}}{s^{\nu+1}},\end{aligned}$$

the interchange of the order of integrations being justified by Fubini's theorem.

Hence by the Parseval relation for the Hankel transform,

$$\begin{aligned}f(s) = \int_0^\infty e^{-st}\, t^{\frac{1}{2}\nu}\, F(t)dt &= \frac{1}{s^{\nu+1}}\int_0^\infty e^{-t/s}\, t^{\frac{1}{2}\nu}\, G(t)dt\\ &= \frac{1}{s^{\nu+1}}\, g\left(\frac{1}{s}\right).\end{aligned}$$

THEOREM 2. *If* $F \in L_2(0, \infty)$, $\nu > -1$,

$$G(x) = \frac{d}{dx}\int_0^\infty k_\nu(xy)\, F(y)\,\frac{dy}{y},$$

where

$$k_\nu(x) = \int_0^x J_\nu(2\sqrt{y})dy,$$

and

$$g(s) = \int_0^\infty e^{-st}\, t^{\frac{1}{2}\nu}\, G(t)dt,$$

then

$$F(t) = \underset{r\to\infty}{\text{l.i.m.}}\; t^{\frac{1}{2}\nu}\, e^{-\frac{1}{2}t}\sum_{n=0}^{r} q_n\,\frac{n!}{\Gamma(\nu+n+1)}\, L_n^{(\nu)}(t)$$

where

$$q_n = (-1)^n 2^{\nu+1}\sum_{r=0}^{n}\binom{n+\nu}{n-r}\frac{4^r}{r!}\, g^{(r)}(2).$$

Proof. By Theorem 1,

$$F(t) = \underset{r\to\infty}{\text{l.i.m.}}\; t^{\frac{1}{2}\nu}\, e^{-\frac{1}{2}t}\sum_{n=0}^{r} q_n\,\frac{n!}{\Gamma(\nu+n+1)}\, L_n^{(\nu)}(t)$$

where

$$q_n = \sum_{r=0}^{n}\binom{n+\nu}{n-r}\frac{1}{r!}\, f^{(r)}\left(\tfrac{1}{2}\right).$$

But in the proof of Lemma 1 we showed that

$$q_n = \text{Residue}_{s=\frac{1}{2}} \left(f(s)\, (s+\tfrac{1}{2})^{n+\nu}/(s-\tfrac{1}{2})^{n+1}\right),$$

and hence using Lemma 2

$$\begin{aligned} q_n &= \text{Residue}_{s=\frac{1}{2}} \left(\frac{1}{s^{\nu+1}}\, g\!\left(\frac{1}{s}\right) (s+\tfrac{1}{2})^{n+\nu}/(s-\tfrac{1}{2})^{n+1}\right) \\ &= \text{Residue}_{s=2} \frac{-1}{2^{\nu-1}} \left(g(s)\, (2+s)^{n+\nu}/(2-s)^{n+1}\right) \\ &= \frac{(-1)^n}{2^{\nu-1}\, n!} \lim_{s\to 2} \frac{d^n}{ds^n} \left(g(s)(2+s)^{n+\nu}\right) \\ &= \frac{(-1)^n}{2^{\nu-1}\, n!} \lim_{s\to 2} \sum_{r=0}^{n} \binom{n}{r} g^{(r)}(s) \frac{\Gamma(n+\nu+1)}{\Gamma(r+\nu+1)} (s+2)^{r+\nu} \\ &= (-1)^n\, 2^{\nu+1} \sum_{r=0}^{n} \binom{n+\nu}{n-r} \frac{4^r}{r!} g^{(r)}(2). \end{aligned}$$

COROLLARY. *Under the hypotheses of Theorem 2, if*

$$f(s) = \int_0^\infty e^{-st}\, t^{\frac{1}{2}\nu}\, F(t)dt,$$

then

$$G(t) = \underset{r\to\infty}{\text{l.i.m.}}\; t^{\frac{1}{2}\nu}\, e^{-\frac{1}{2}t} \sum_{n=0}^{r} q'_n \frac{n!}{\Gamma(\nu+n+1)} L_n^{(\nu)}(t)$$

where

$$q'_n = (-1)^n\, 2^{\nu+1} \sum_{r=0}^{n} \binom{n+\nu}{n-r} \frac{4^r}{r!} f^{(r)}(2).$$

Proof. This follows from Theorem 2 since the relation between F and G is reciprocal.

REFERENCES

1. G. Doetsch, *Handbuch der Laplace Transformation* I (Basel, 1950).
2. A. Erdélyi *et al.*, *Higher Transcendental Functions* II (New York, 1953).
3. A. Erdélyi *et al.*, *Tables of Integral Transforms* I (New York, 1954).
4. P. G. Rooney, *On an inversion formula for the Laplace transformation*, Can. J. Math., *7* (1955), 101–115.
5. J. Shohat, *Laguerre polynomials and the Laplace transform*, Duke Math. J., *6* (1940), 615–626.
6. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford, 1948).
7. D. V. Widder, *The Laplace Transform* (Princeton, 1941).

University of Toronto

ON SOME RELATIONS BETWEEN PARTIAL AND ORDINARY DIFFERENTIAL EQUATIONS

ERWIN KREYSZIG*

Dedicated to Professor Dr. A. WALTHER on his 60th birthday.

1. Introduction. The theory of solutions of partial differential equations

$$\Delta u + \alpha(x, y)u_x + \beta(x, y)u_y + \gamma(x, y)u = 0 \tag{1.1}$$

with analytic coefficients can be based upon the theory of analytic functions of a complex variable; the basic tool in this approach is integral operators which map the set of solutions of (1.1) onto the algebra of analytic functions. For certain classes of operators this mapping which is first defined in the small, can be continued to the large, cf. Bergman **(3)**. In this way theorems on analytic functions give rise to theorems on (real and complex) solutions of (1.1). Some of the operators possess a remarkable property: they generate solutions of certain partial differential equations (1.1) which also satisfy ordinary linear differential equations in x or y. This was first observed by Bergman **(1;2)** in the special case of the equation $\Delta u + u = 0$. This property is of interest since it permits the investigation of such solutions of (1.1) by means of the theory of ordinary differential equations. The present paper is concerned with a class of partial differential equations (1.1) which possess solutions of that type. We shall derive an infinite set of independent particular solutions and obtain relations between singularities of the coefficients of (1.1) and those of the corresponding ordinary differential equations; cf. §§ 3–5. These results will enable us to characterize some basic properties of those solutions of (1.1); cf. § 6.

2. Partial differential equations of class $\mathfrak{C}$. If we introduce the variables $z = x + iy$, $z^* = x - iy$, the equation (1.1) takes the form

$$u_{zz^*} + a(z, z^*)u_z + b(z, z^*)u_{z^*} + c(z, z^*)u = 0 \tag{2.1}$$

where

$$u_{zz^*} = \tfrac{1}{4}\Delta u, \quad u_z = \tfrac{1}{2}(u_x - iu_y), \quad u_{z^*} = \tfrac{1}{2}(u_x + iu_y),$$
$$a = \tfrac{1}{4}(\alpha + i\beta), \quad b = \tfrac{1}{4}(\alpha - i\beta), \quad c = \tfrac{1}{4}\gamma.$$

If we set

$$u = U \exp\left(-\int_0^{z^*} a(z, t)dt\right)$$

we obtain from (2.1)

*Received Oct. 21, 1957. This work was done while the author attended the Summer Research Institute of the Canadian Mathematical Congress, 1956. The author wishes to thank Professor R. L. Jeffery (Queen's University, Kingston, Ont.) for his interest and help.

$$L(U) \equiv U_{zz^*} + B(z, z^*)U_{z^*} + C(z, z^*)U = 0 \tag{2.2}$$

where

$$B = b - \int_0^{z^*} a_z(z, t)dt, \quad C = c - a_z - ab.$$

We note that for complex values of x and y the variables z and z^* are independent.

Definition 1. *An operator of the form*

$$U(z, z^*) \equiv P(f) = \int_{-1}^{1} E(z, z^*, t)f(\tfrac{1}{2}z(1 - t^2))(1 - t^2)^{-\frac{1}{2}}dt \tag{2.3}$$

is called a Bergman operator. In (2.3) *the "associated function"* $f(z)$ *of* $U(z,z^*)$ *is an analytic function of a complex variable regular at the origin. The "generating function"* $E(z,z^*,t)$ *is independent of the special choice of* $f(z)$.

In order that $U(z,z^*)$ be a solution of (2.2) the function $E(z,z^*,t)$ must satisfy the equation

$$(1 - t^2)E_{z^*t} - t^{-1}E_{z^*} + 2zt\, L(E) = 0, \tag{2.4}$$

as can be seen by inserting (2.3) into (2.2).

Definition 2. *A partial differential equation* (2.2) *is said to be of the class* $\mathfrak{E}$ *if its solutions can be generated in the form* (2.3) *with a generating function of the type*

$$E(z, z^*, t) = \exp Q(z, z^*, t), \qquad Q(z, z^*, t) = \sum_{\mu=1}^{m} q_\mu(z, z^*)t^\mu. \tag{2.5}$$

Necessary and sufficient conditions have been obtained for the coefficients of (2.2) in order that (2.2) should be of the class $\mathfrak{E}$; cf. Kreyszig **(4)**.

3. Existence of ordinary differential equations satisfied by solutions of partial differential equations of the class $\mathfrak{E}$. If in (2.3), $f(z) = z^n$, $n = 0, 1, \ldots$, the corresponding solutions of the partial differential equations of the class $\mathfrak{E}$ satisfy a linear ordinary differential equation; cf. Kreyszig **(5)**. It was conjectured that the (more important) solutions with meromorphic associated functions have a similar property. However, the method used in **(5)** fails in this case. In order to treat this problem in a systematic way we first consider solutions $U(z,z^*)$ which correspond to associated functions

$$f_n(z) = (z - \zeta)^{-n}, \qquad \zeta \neq 0, \qquad n = 1, 2, \ldots. \tag{3.1}$$

In order to derive ordinary differential equations satisfied by $U(z,z^*)$ we have to consider this function in certain planes of the (real four-dimensional) zz^*-space. The form of these equations will depend on the choice of these planes. We take the planes $y = y_0 = $ const. Then we have the advantage htat $U(z,z^*) \equiv \tilde{U}(x,y)$ is an analytic function of x.

THEOREM 1. *Each solution* $\tilde{U}(x,y) = U(z,z^*)$ *of a partial differential equation* (2.2) *of the class* $\mathfrak{E}$ *with an associated function* (3.1) *satisfies an ordinary linear differential equation*

$$N(U) \equiv \tilde{N}(\tilde{U}) = \sum_{\rho=0}^{r} G_\rho(x, y_0) \frac{d^\rho \tilde{U}}{dx^\rho} = 0, \qquad G_r = 1, y = y_0 = \text{const}, \tag{3.2}$$

of order

$$r < m + 3. \tag{3.3}$$

The coefficients $G_\rho(x,y) \equiv g_\rho(z,z^*)$ *are rational functions of* $q_\mu(z,z^*)$, $\mu = 0,1, \ldots, m$. *The order* r *is independent of* n.

Proof. In consequence of (2.5) and (3.1) the integrand of (2.3) takes the form

$$\begin{aligned} J(x, y, t) &\equiv j(z, z^*, t) = \exp Q(z, z^*, t)\, s(z, t)^{-n}(1 - t^2)^{-\frac{1}{2}}, \\ S(x, y, t) &\equiv s(z, t) = \tfrac{1}{2}z(1 - t^2) - \zeta. \end{aligned} \tag{3.4}$$

It suffices to prove that J satisfies the non-homogeneous equation

$$\tilde{N}(J) = R, \quad R(x, y, t) = \frac{d}{dt}[(1 - t^2)H(x, y, t)] \tag{3.5}$$

where H is a regular function of t for $|t| < 1$. If we integrate both sides of this equation with respect to t from -1 to 1 we obtain (3.2). We choose

$$H = P\, S^{-r+1} J \tag{3.6}$$

where

$$P(x, y, t) \equiv p(z, z^*, t) = \sum_{\lambda=0}^{l} p_\lambda(z, z^*)\, t^\lambda; \tag{3.7}$$

the degree l and the coefficients $p_\lambda(z,z^*)$ will be suitably determined, see below. We have

$$J_t \equiv \frac{\partial J}{\partial t} = (Q_t + (1 - t^2)^{-1}t + s^{-1}nzt)\, J, \tag{3.8}$$

$$J^{(\alpha)} \equiv \frac{\partial^\alpha J}{\partial x^\alpha} = T_\alpha\, J \tag{3.9}$$

where

$$T_1 = \frac{\partial Q}{\partial x} - n(1 - t^2)(2S)^{-1} \tag{3.10}$$

and

$$T_\alpha = |A_{\alpha 1}\, A_{\alpha 2} \ldots A_{\alpha\alpha}|\,, \qquad \alpha = 2, 3, \ldots, \tag{3.11}$$

is a determinant with the column vectors

$$A_{\alpha,\beta+1} = \left(\binom{\beta}{\beta} T_1^{(\beta)}, \binom{\beta}{\beta - 1} T_1^{(\beta-1)}, \ldots, \binom{\beta}{0} T_1, -1, 0, 0, \ldots\right),$$

$$T_1^{(\beta)} \equiv \frac{\partial^\beta T_1}{\partial x^\beta}, \qquad \beta = 0, 1, \ldots, \alpha - 1;$$

(the number of zeros decreases with increasing β; in $A_{\alpha,\alpha-1}$ there are no more zeros left, and in $A_{\alpha\alpha}$ the term T_1 is the last one). This can easily be proved by induction. From (3.5) – (3.8) we find

$$(3.12)\qquad R = \{ -tP + (1-t^2)(P_t + P[S^{-1}(r+n-1)tz + Q_t])\}\ J\, S^{-r+1}.$$

If we insert (3.9) and (3.12) into (3.5), omit the common factor J and multiply each term by S^r, each side of the resulting equation becomes a polynomial in t. If we choose

$$(3.13)\qquad l = (m+2)r - m - 3,$$

cf. (3.7), these two polynomials have the same degree, namely $(m+2)r$. In the equation thus obtained the coefficients of each power of t must be the same on both sides. Hence we obtain a system of $(m+2)r+1$ linear equations. If we choose

$$(3.14)\qquad r = m+3$$

the number of equations equals the total number of the coefficients $G_0, \ldots, G_{r-1}$ of (3.2) and of the coefficients $p_0, \ldots, p_l$ of (3.7). In order to be able to determine these functions G_ρ and p_λ it suffices that the determinant $D(z,z^*)$ of the coefficients of the system does not vanish identically, since every neighbourhood of a point of a zero surface of D contains always points at which $D(z,z^*) \not\equiv 0$. Furthermore, it can readily be seen that the rank of D is always different from zero. Hence if $D(z,z^*) \equiv 0$ there exists a subdeterminant of D which does not vanish identically. In the case $D \equiv 0$ the order r of (3.2) reduces to values smaller than $m+3$; cf. (2.5), and the coefficients of (3.2) can be determined in a similar manner. This completes the proof.

This result may be extended to the case of solutions with arbitrary rational associated functions as follows.

THEOREM 2. *Each solution $\tilde{U}(x,y) \equiv U(z,z^*)$ of a partial differential equation* (2.2) *of the class $\mathfrak{E}$ with a rational associated function $f(z)$ satisfies an ordinary linear differential equation in x whose coefficients are rational functions of $q_0, \ldots, q_m$, cf.* (2.5). *If $f(z)$ has poles of orders β_κ at $z = z_\kappa$, $\kappa = 1,2,\ldots,k$, the equation has the order*

$$(3.15)\qquad r \leqslant (\alpha+\beta+1)m + \alpha + 3\beta$$

where $\beta = \beta_1 + \beta_2 + \ldots + \beta_k$ and α is the degree of the polynomial $F_1(z)$ in the representation of $f(z)$ as a sum of $F_1(z)$ and a proper rational function $F_2(z)$; m is defined by (2.5).

Proof. The polynomial $F_1(z)$ is a sum of at most $\alpha+1$ terms. To each of these terms and to each partial fraction of $F_2(z)$ there corresponds a particular solution $\tilde{U}_\delta(x,y)$ of (2.2). We thus have

$$(3.16)\qquad \tilde{U}(x,y) = \sum_{\delta=1}^{d} \tilde{U}_\delta(x,y), \qquad d \leqslant \alpha+\beta+1.$$

Each of the functions $\tilde{U}_s(x,y)$ corresponding to $F_1(z)$ satisfies an ordinary linear differential equation of the order $r^* < m + 1$, cf. (**5**, Theorem 2), while each of the other functions satisfies such an equation of the order $r^{**} < m + 3$, cf. Theorem 1 of this paper. Thus, we have a system (S) of ordinary linear differential equations whose coefficients are rational functions of $q_0, \ldots, q_m$, cf. (2.5). We differentiate each of these differential equations and also the equation (3.16) r times and eliminate all the functions $\tilde{U}_s(x,y)$ and their derivatives from the enlarged system (S^*) thus obtained. In order to be able to do so we have to choose r so that the number of equations of (S^*) equals the number of functions to be eliminated. It can easily be seen that r cannot be greater than $(\alpha + \beta + 1)\, m + \alpha + 3\beta$. Since we differentiated (3.16) r times the r^{th} derivative of $\tilde{U}(x,y)$ is the highest one which occurs in (S^*). This completes the proof.

4. Subclasses of the class $\mathfrak{E}$. The coefficients $B(z,z^*)$ and $C(z,z^*)$ of the partial differential equations (2.2) of the class $\mathfrak{E}$ are related to the coefficients $q_\nu(z,z^*)$ of the generating function (2.5) as follows (**4**, Theorem 1).

(I) *If* $q_1(z,z^*) \not\equiv 0$ *then*

$$B = -\frac{\partial q_0}{\partial z} - \frac{q_2}{z}, \quad C = -\frac{q_1}{2z}\frac{\partial q_1}{\partial z^*}. \tag{4.1}$$

(II) *If* $q_1 \equiv 0$ *then also* $q_3 \equiv 0,\ q_5 \equiv 0 ,\ldots,$ *and*

$$B = -\frac{\partial q_0}{\partial z} - \frac{q_2}{z}, \quad C = -\frac{1}{2z}\frac{\partial q_2}{\partial z^*}; \tag{4.2}$$

q_0 depends only on z and can have singularities. In case (I) q_1 depends on z and z^* and can have singularities, considered as a function of z^* for any finite constant value of z. In case (I), q_2 is regular while in case (II) q_2, considered as a function of z^* for any finite constant value of z, can have singularities.

Hence the class $\mathfrak{E}$ consists of two subclasses $\mathfrak{E}_{\text{I}}$ and $\mathfrak{E}_{\text{II}}$ corresponding to the two cases (I) and (II).

In case (II) the function $Q(z,z^*,t)$, defined by (2.5), is an even function of t. Hence, in this case, the functions T_n, cf. (3.10), (3.11), are also even functions of t. Let $P(x,y,t)$ be an odd function of t; then $R\,S^r J^{-1}$ is an even function of t; cf. (3.4)–(3.7). Hence, in this case the polynomials considered in the proof of Theorem 1 are even functions of t and have the degree $(m + 2)r$. The function $P(x,y,t)$ has now only $\frac{1}{2}(l + 1)$ coefficients $p_\lambda(z,z^*)$ where l is defined by (3.13). The total number $\frac{1}{2}(l + 1) + r$ of the functions G_ρ and p_λ must equal the number of powers occurring in the above-mentioned polynomials. We thus obtain the result that each solution of a partial differential equation (2.2) of the subclass $\mathfrak{E}_{\text{II}}$ with an associated function (3.1) satisfies an ordinary linear differential equation of the order

$$r = < \tfrac{1}{2}m + 2, \qquad (m \text{ even}). \tag{4.3}$$

It can be similarly proved that such a solution with an associated function $f_n(z) = z^n$, $n = 0,1,\ldots,$ satisfies an ordinary linear differential equation of the order

$$r < \tfrac{1}{2}\,m + 1, \qquad (m \text{ even}). \tag{4.4}$$

Applying to these results the idea of the proof of Theorem 2 we obtain the following

COROLLARY. *Each solution of a partial differential equation* (2.2) *of the subclass* $\mathfrak{C}_{\mathrm{II}}$ *with a rational associated function satisfies an ordinary differential equation in* x *of the order*

$$r < m + \beta + (\tfrac{1}{2}m + 1)(\alpha + \beta), \qquad (m \text{ even}), \tag{4.5}$$

where α *and* β *are defined as in Theorem 2. The coefficients of this equation are rational functions of* $q_0, \ldots, q_m$, cf. (2.5).

Partial differential equations of the subclass $\mathfrak{C}_{\mathrm{II}}$ thus have the remarkable property that the corresponding ordinary differential equations have a smaller order than those corresponding to partial differential equations of the subclass $\mathfrak{C}_{\mathrm{I}}$.

5. Relations between singularities of the partial differential equation (2.2) **and those of the corresponding ordinary differential equation** (3.2). The relations between $q_0, \ldots, q_m$ (cf. (2.5)) and the coefficients B, C of (2.2) on the one hand, and between $q_0, \ldots, q_m$ and the coefficients G_ρ of (3.2) on the other hand, enable us to obtain direct relations between the singularities of the given partial differential equation (2.2) and the ordinary differential equation (3.2) which we have derived. Since the procedure of obtaining such relations is similar to that developed in **(5)** we omit details and state the result only. We find

THEOREM 3. *The singularities of the ordinary differential equation* (3.2) *and those of the corresponding partial differential equation* (2.2) *of the class* $\mathfrak{C}$ *are related as follows.*

(*i*) *If* B, *considered as a function of* z *for any finite value* $z^* = const$, *has a pole of the order* s *at a point* $z = a$, *the coefficient* G_ρ *of* (3.2), *considered as a function of* z, *has a pole of the order*

$$s_1(s,\rho) = sw_\rho, \qquad w_\rho = m + 3 - \rho, \tag{5.1}$$

at $z = a$. *If* $s = 1$ *then* (3.2) *is of Fuchsian type at* $z = a$.

(*ii*) *If* $q_1(z,z^*) \not\equiv 0$ *and* C, *considered as a function of* z^* *for any finite value* $z = const$, *has a pole of the order* $2s-1$, $s > 1$, *at a point* $z^* = a^*$, *the coefficient* G_ρ *of* (3.2), *considered as a function of* z^*, *has a pole of the order*

$$s_2(s,\rho) = \begin{cases} s\,w_\rho - (s-1)\epsilon_\rho & (m = 1) \\ s\,w_\rho + (s-1)\epsilon_{\rho+1} & (m > 2) \end{cases} \tag{5.2}$$

at $z^* = a^*$, where

$$\epsilon_\alpha = \begin{cases} 0 & (\alpha \text{ even}) \\ 1 & (\alpha \text{ odd}). \end{cases}$$

It should be noted that, for a fixed value of m, these relations are the same for all solutions of (2.2) with the associated functions (3.1).

6. Final remark. Let us finally state some remarks about the characterization of solutions of (2.2) by means of the preceding results.

(a) The solutions $U(z,z^*) \equiv \tilde{U}(x,y)$ of partial differential equations (2.2) of the class $\mathfrak{E}$ with rational associated functions also satisfy an ordinary linear differential equation, *considered as functions of y for any finite value $x = x_0 = const$*, as can be proved by using the preceding methods. This result and the results obtained in §§ 3–5 enable us to investigate these (single or multi-valued) solutions of (2.2) outside of the domain of validity of the integral representation (2.3). An appropriate theory of this kind **(2)** leads to a characterization of the behaviour of the solutions in the neighbourhood of branch surfaces and some other basic properties; the theory can immediately be applied to the class of equations (2.2) under consideration, but we should stress the fact that for this purpose we need the detailed information about the ordinary differential equations which is given by the preceding theorems.

(b) The coefficients of the ordinary differential equations satisfied by $U(z,z^*) \equiv \tilde{U}(x,y)$ are rational functions of $q_0, \ldots, q_m$. In the special case of partial differential equations (2.2) with rational coefficients the coefficients of the ordinary differential equations are *rational functions of x and y*, respectively. Hence, in this case, the singularities of the solutions of (2.2) with rational associated functions lie on two-dimensional algebraic manifolds in the real four-dimensional space.

(c) So far we have obtained conditions on the associated functions of the solutions $U(z,z^*) \equiv \tilde{U}(x,y)$ of (2.2) in order that $U(x,y)$ satisfies ordinary differential equations. These conditions may be replaced by conditions on the coefficients $a_{\kappa\lambda}$ of the development

$$U(z,z^*) = \sum_{\kappa,\lambda=0}^{\infty} a_{\kappa\lambda} z^{\kappa} z^{*\lambda}. \tag{6.1}$$

Let the associated function $f(z)$ of $U(z,z^*)$ be represented in the form

$$f(z) = \sum_{\nu=0}^{\infty} c_\nu z^\nu \tag{6.2}$$

and the generating function (2.5) of the operator (2.3) in the form

$$E(z, z^*, t) = \exp Q(z, z^*, t) = \sum_{\mu,\sigma=0}^{\infty} b_{\mu\sigma}(t)\, z^{\mu} z^{*\sigma}. \tag{6.3}$$

Then, by (2.3),

$$U(z, 0) = \sum_{\kappa=0}^{\infty} a_{\kappa 0} z^{\kappa} = \sum_{\mu,\nu=0}^{\infty} c_{\nu} A_{\mu\nu} z^{\mu+\nu}$$

where

$$A_{\mu\nu} = 2^{-\nu} \int_{-1}^{1} b_{\mu 0}(t)(1 - t^2)^{\nu-\frac{1}{2}} dt.$$

By comparing the coefficients of corresponding powers of z on both sides we obtain

$$a_{\kappa 0} = \sum_{\nu=0}^{\kappa} c_{\nu} A_{\kappa-\nu,\nu}\,, \qquad \kappa = 0, 1, \ldots . \tag{6.4}$$

The solution of this system yields representations of the coefficients c_ν of the associated function in terms of the coefficients $a_{\kappa 0}$ of the development (6.1). Using these representations and theorems by Hadamard **(6)** we obtain information on the nature and location of the singularities of the associated function of $U(z,z^*)$ from the sequence $\{a_{\kappa 0}\}$ of the coefficients in (6.1). This yields sufficient conditions on the coefficients $\alpha_{\kappa 0}$ in order that $U(z,z^*) \equiv \tilde{U}(x,y)$ satisfy ordinary linear differential equations. In this connection the important problem arises as to what extent similar conclusions can be drawn if other subsequences, say $\{a_{\kappa\lambda}\}$, $\lambda > 0$ and fixed, of the coefficients in (6.1) are known. This question will be considered in another paper.

References

1. S. Bergman, *Ueber Kurvenintegrale von Funktionen zweier komplexer Veraenderlicher, die die Differentialgleichung $\Delta V + V = 0$ befriedigen*, Math. Z. *32* (1930), 386–405.

2. ———, *Zur Theorie der Funktionen, die eine lineare partielle Differentialgleichung befriedigen*, Rec. Math. (Mat. Sbornik) N.S. *2* (1937), 1169–1198.

3. ———, *Linear operators in the theory of partial differential equations*, Trans. Amer. Math. Soc. *53* (1943), 130–155.

4. E. Kreyszig, *On a class of partial differential equations*, J. Rat. Mech. Anal. *4* (1955), 907–923.

5. ———, *On certain partial differential equations and their singularities*, J. Rat. Mech. Anal. *5* (1956), 805–820.

6. J. Hadamard, *Essai sur l'étude des fonctions données par leur développement de Taylor*, J. de Math. pur. appl. (4) *8* (1892), 101–186.

University of Ottawa
and
Ohio State University

A NEW PROOF OF A THEOREM OF LEHMER

J. B. ROBERTS

In 1851 Prouhet **(2)** stated that any b^{k+1} consecutive positive integers (b a positive integer $\geqslant 2$) can be separated into b sets $\bar{C}_0, \ldots, \bar{C}_{b-1}$ each with b^k members in such a way that

$$\sigma_t(\bar{C}_0) = \ldots = \sigma_t(\bar{C}_{b-1}), \qquad 0 \leqslant t \leqslant k,$$

where $\sigma_t(\bar{C}_j)$ designates the sum of the tth powers of the numbers in $\bar{C}_j$. In 1947 Lehmer **(1)** generalized this result.

This paper contains a new proof of Lehmer's theorem. The proof gives a slightly more general result which is immediate from the less general form. The method of proof is similar to that of Lehmer except that it makes use of difference operators rather than differentiation. Two other proofs of Lehmer's theorem which are based on rather different ideas were given by Wright **(3)**.

1. Lehmer's theorem.

THEOREM. *Let $\alpha_0, \ldots, \alpha_k$ be an arbitrary set of $k + 1$ complex numbers (distinct or not) and let b be an integer $\geqslant 2$. Let C be the collection of all numbers of the form $j_0\alpha_0 + \ldots + j_k\alpha_k$ where the j_t are integers satisfying $0 \leqslant j_t \leqslant b - 1$. Further, let C_j, $0 \leqslant j \leqslant b - 1$, be the collection of elements of C for which $j_0 + \ldots + j_k \equiv j$ (mod b). Then for $P(x)$, a complex polynomial of degree smaller than or equal to k,*

$$\sum_{n \epsilon C_i} P(x + n) = \sum_{n \epsilon C_j} P(x + n), 0 \leqslant i \leqslant b - 1, 0 \leqslant j \leqslant b - 1.$$

2. Proof of the theorem via operators. Let $E(c)$, where c is an arbitrary complex number, be an operator which maps the (complex) polynomial $P(x)$ onto the polynomial $P(x + c)$. Then $E(a + c) = E(a)E(c)$.

Throughout the remainder of this paper b is to be a fixed integer $\geqslant 2$ and ω any bth root of unity.

LEMMA 1.

$$\prod_{m=0}^{k} \sum_{j=0}^{b-1} \omega^j E(j\alpha_m) = \sum_{n \epsilon C} \omega^{v(n)} E(n),$$

where $v(n) = j_0 + \ldots + j_k$ when $n = j_0\alpha_0 + \ldots + j_k\alpha_k$. Further, if a number $n \in C$ has more than one representation of the specified kind the right-hand sum includes a term for each representation.

Received May 6, 1957.

The proof of this lemma is immediate upon multiplying out the left side of the equation.

If we now take $\omega \neq 1$ we find

$$\sum_{j=0}^{b-1} \omega^j E(j\alpha_m)P(x) = \sum_{j=0}^{b-1} \omega^j P(x + j\alpha_m) = a_q x^q \sum_{j=0}^{b-1} \omega^j + \sum_{j=0}^{b-1} \omega^j Q(x, j)$$

where a_q is the leading coefficient of $P(x)$ and $Q(x, j)$ is a polynomial of degree smaller than that of $P(x)$. Now, since

$$\sum_{j=0}^{b-1} \omega^j = 0$$

we have

Lemma 2.

$$\sum_{j=0}^{b-1} \omega^j E(j\alpha_m), \omega \neq 1,$$

maps a polynomial $P(x)$ onto a polynomial of smaller degree.

An immediate consequence of this lemma is that when $P(x)$ has degree smaller than or equal to k the operator on the left side of the equation in Lemma 1 maps $P(x)$ onto 0. Hence

$$(1) \quad \sum_{n \in C} \omega^{s(n)} E(n)P(x) = \sum_{n \in C} \omega^{s(n)} P(x + n) = \sum_{j=0}^{b-1} \omega^j \sum_{n \in C_j} P(x + n) = 0.$$

Equation (1) holds for all bth roots of unity other than 1. But when $c_{b-1}x^{b-1} + \ldots + c_0 = 0$ for all bth roots of unity other than 1 we must have

$$\begin{aligned} c_{b-1}x^{b-1} + \ldots + c_0 = c_{b-1}(x - \omega_1) \ldots (x - \omega_{b-1}) &= c_{b-1}(x^{b-1} + \ldots + 1) \\ &= c_{b-1}x^{b-1} + \ldots + c_{b-1} \end{aligned}$$

for all x. Hence $c_i = c_{b-1}$ for $0 < i < b - 1$. ($\omega_1, \ldots, \omega_{b-1}$ are the bth roots of unity other than 1.) Applying this result to (1) we find that

$$\sum_{n \in C_j} P(x + n)$$

is independent of j. This completes the proof of Lehmer's theorem.

3. Special Cases.

(a) If we take $\alpha_i = b^i$, $0 < i < k$, in the theorem then C_j, $0 < j < b - 1$, consists of those integers from 0 to $b^{k+1} - 1$ whose base b digit sum is congruent to j modulo b. Hence, if we take $P(x) = x^t$ and put $x = a + 1$, $a > 0$, the theorem yields

$$\sigma_t(\bar{C}_0) = \ldots = \sigma_t(\bar{C}_{b-1}), \qquad 0 < t < k,$$

where $\bar{C}_j$ consists of those integers m from $a + 1$ to $a + b^{k+1}$ such that $m - (a + 1) \in C_j$. This is Prouhet's result applied to the b^{k+1} consecutive integers $a + 1, \ldots, a + b^{k+1}$.

(b) Take the α_i as in (a) and let $P(x) = (j + mx) \ldots (j + mx - q + 1)/q!$ When one puts $x = p/q$, in this case the theorem yields the result that

$$\sum_{n \in C_j} \binom{j + q + mn}{q}$$

is independent of j. The C_j are as in (a).

4. Calculation of the C_j when $\alpha_i = b^i$. We illustrate the calculation by means of an example. We take $b = 3$, $k = 2$. The aim is to construct a string of twenty-seven symbols of three kinds which, when attached to the integers 0 through twenty-six, have the property that two of these integers are in the same C_j, $0 < j < 2$ in this case, if and only if they have the same attached symbol. We use α, β, γ as our three kinds of symbols. The construction proceeds as follows.

$$\begin{gathered} \alpha\beta\gamma \\ \alpha\beta\gamma\ \beta\gamma\alpha\ \gamma\alpha\beta \\ \alpha\beta\gamma\ \beta\gamma\alpha\ \gamma\alpha\beta\ \beta\gamma\alpha\ \gamma\alpha\beta\ \alpha\beta\gamma\ \gamma\alpha\beta\ \alpha\beta\gamma\ \beta\gamma\alpha \end{gathered}$$

If k had been 3 we would have continued one more step to get a string of 81 digits, the first twenty-seven of which would have been those in the third line above, the next twenty-seven of which would have been the cyclic permutation of those above beginning with the second block of 9, and the last twenty-seven would have been the cyclic permutation of those above beginning with the last block of 9.

In our case we have

$$\begin{array}{ccccccccccccccccccccc} \alpha & \beta & \gamma & \beta & \gamma & \alpha & \gamma & \alpha & \beta & \beta & \gamma & \alpha & \gamma & \alpha & \beta & \alpha & \beta & \gamma & \gamma & \alpha & \beta \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \end{array}$$

$$\begin{array}{cccccc} \alpha & \beta & \gamma & \beta & \gamma & \alpha \\ 21 & 22 & 23 & 24 & 25 & 26 \end{array}$$

and therefore

$$\begin{aligned} C_0 &= \{0, 5, 7, 11, 13, 15, 19, 21, 26\} \\ C_1 &= \{1, 3, 8, 9, 14, 16, 20, 22, 24\} \\ C_2 &= \{2, 4, 6, 10, 12, 17, 18, 23, 25\}. \end{aligned}$$

Hence, to split the 27 positive integers $r, r + 1, \ldots, r + 26$ into three classes satisfying Prouhet's result we take the a of section 3(a) to be $r - 1$ and find

$$\begin{aligned} \bar{C}_0 &= \{r, r + 5, r + 7, r + 11, r + 13, r + 15, r + 19, r + 21, r + 26\} \\ \bar{C}_1 &= \{r + 1, r + 3, r + 8, r + 9, r + 14, r + 16, r + 20, r + 22, r + 24\} \\ \bar{C}_2 &= \{r + 2, r + 4, r + 6, r + 10, r + 12, r + 17, r + 18, r + 23, r + 25\}. \end{aligned}$$

If $r = 1$, for instance, we have

$$\begin{aligned} &1^t + 6^t + 8^t + 12^t + 14^t + 16^t + 20^t + 22^t + 27^t \\ &= 2^t + 4^t + 9^t + 10^t + 15^t + 17^t + 21^t + 23^t + 25^t \\ &= 3^t + 5^t + 7^t + 11^t + 13^t + 18^t + 19^t + 24^t + 26^t, \qquad 0 < t < 2. \end{aligned}$$

References

1. D. H. Lehmer, *The Tarry-Escott problem*, Scripta Math., *13* (1947), 37–41.
2. M. E. Prouhet, *Memoire sur quelques relations entre les puissances des nombres*, C.R. Acad. Sci., *33* (Paris, 1851), 225.
3. E. M. Wright, *Equal sums of like powers*, Proc. Edin. Math. Soc., (2), *8* (1949), 138–142.

Wesleyan University
Middletown, Conn.
and
Reed College
Portland, Oregon

CERTAIN TRANSFORMATIONS OF NEARLY-POISED BILATERAL HYPERGEOMETRIC SERIES OF SPECIAL TYPE

H. S. SHUKLA

1. Introduction. A few years ago Bailey **(1)** gave certain transformations of both terminating and non-terminating nearly-poised hypergeometric series of the ordinary type and later on he also deduced basic analogues of some of his transformations. Recently, **(3)** I gave certain transformations of both ordinary and basic terminating nearly-poised bilateral hypergeometric series which generalized Bailey's results. Since transformations of nearly-poised series have not been systematically studied so far, I deduced in another paper **(4)** certain relations of both ordinary and basic bilateral series which involved either only nearly-poised series or both terminating well-poised and non-terminating nearly-poised series. In this paper I obtain certain transformations of non-terminating nearly-poised bilateral series of special types ${}_4H_4$ and ${}_5H_5$ and these transformations are generalizations of Bailey's known results. In the sequel the sum of a particular ${}_3H_3$ is also given and is believed to be new.

The following notation is used throughout the paper:

$$(a)_n = a(a+1)\dots(a+n-1);\ (a)_0 = 1;\ (a)_{-n} = (-1)^n/(1-a)n;$$

$${}_rH_r\left[\begin{matrix} a_1, a_2, \dots, a_r; z \\ b_1, b_2, \dots, b_r \end{matrix}\right] = \sum_{n=-\infty}^{\infty} \frac{(a_1)_n (a_2)_n \dots (a_r)_n}{(b_1)_n (b_2)_n \dots (b_r)_n} z^n;$$

$$\Gamma\left[\begin{matrix} a_1, a_2, \dots, a_r; \\ b_1, b_2, \dots, b_r \end{matrix}\right] = \frac{\Gamma(a_1)\Gamma(a_2)\dots\Gamma(a_r)}{\Gamma(b_1)\Gamma(b_2)\dots\Gamma(b_r)}.$$

Also, idem $(a; b)$ means that the preceding expression is repeated with a and b interchanged.

2. In a recent paper **(4)** I have deduced the following relation between M nearly-poised hypergeometric series of the type ${}_MH_M$ with unit argument:

$$\Gamma\left[\begin{matrix} a_2, a_3, \dots, a_M, 1-a_2, 1-a_3, \dots, 1-a_M; \\ b_1, c_1+b_1-c_2, c_1+b_1-c_3, \dots, c_1+b_1-c_{M-1}, b_M, 1-c_1, \dots, 1-c_M \end{matrix}\right]$$

$$\times\ {}_MH_M\left[\begin{matrix} c_1, c_2, \dots, c_{M-1}, c_M; \\ b_1, c_1+b_1-c_2, \dots, c_1+b_1-c_{M-1}, b_M \end{matrix}\right]$$

$$(2.1)\quad +\ \Gamma\left[\begin{matrix} a_2-1, 2-a_2, a_2-a_3, \dots, a_2-a_M, 1+a_3-a_2, \dots, \\ 1+a_M-a_2; \\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, \dots, 1+c_1+b_1-c_{M-1} \\ -a_2, 1+b_M-a_2, a_2-c_1, \dots, a_2-c_M \end{matrix}\right]$$

Received October 1, 1957.

$$\times\, {}_MH_M\begin{bmatrix} 1+c_1-a_2, 1+c_2-a_2, \ldots, 1+c_{M-1}-a_2, 1+c_M-a_2; \\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, \ldots, \\ 1+c_1+b_1-c_{M-1}-a_2, 1+b_M-a_2 \end{bmatrix}$$
$$+ \text{idem}\,(a_2; a_3, a_4, \ldots, a_M) = 0.$$

The transformation (2.1) can be deduced directly from Slater's transformation (**2**, (10)).

If we take $M = 4$ in (2.1) and then reverse the first ${}_4H_4$ series on the left and put $c_4 = 0$ we get the following relation between a nearly-poised ${}_4F_3$ series of the first kind and three nearly-poised ${}_4H_4$ series:

$$\Gamma\begin{bmatrix} a_2, a_3, a_4, 1-a_2, 1-a_3, 1-a_4; \\ b_1, c_1+b_1-c_2, c_1+b_1-c_3, b_4, 1-c_1, 1-c_2, 1-c_3 \end{bmatrix}$$
$$\times\, {}_4F_3\begin{bmatrix} 1-b_4, 1-b_1, 1+c_2-c_1-b_1, 1+c_3-c_1-b_1; \\ 1-c_1, \quad 1-c_2, \quad 1-c_3 \end{bmatrix}$$
$$(2.2) \quad +\Gamma\begin{bmatrix} a_2-1, 2-a_2, 1+a_3-a_2, 1+a_4-a_2, a_2-a_3, a_2-a_4; \\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, 1+c_1+b_1-c_3-a_2, \\ 1+b_4-a_2, a_2-c_1, a_2-c_2, a_2-c_3, a_2 \end{bmatrix}$$
$$\times\, {}_4H_4\begin{bmatrix} 1+c_1-a_2, 1+c_2-a_2, 1+c_3-a_2, 1-a_2; \\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, 1+c_1+b_1-c_3-a_2, \\ 1+b_4-a_2 \end{bmatrix}$$
$$+ \text{idem}\,(a_2; a_3, a_4) = 0.$$

Since the nearly-poised ${}_4F_3$ series of the first kind on the left of (2.2) can be expressed in terms of two Saalschützian ${}_5F_4$ series (**1**, § 6.5 (1)), we get the following relation between two Saalschützian ${}_5F_4$ and three nearly-poised ${}_4H_4$ series:

$$\Gamma\begin{bmatrix} a_2, a_3, a_4, 1-a_2, 1-a_3, 1-a_4, 2b_1+c_1-c_2-c_3-1; \\ b_1, c_1+b_1-c_2, c_1+b_1-c_3, b_4, c_1+b_1-c_2-c_3, b_1-c_2, b_1-c_3, \\ 2-b_1-c_1 \end{bmatrix}$$
$$\times\, {}_5F_4\begin{bmatrix} 1-b_1, 1+c_2-c_1-b_1, 1+c_3-c_1-b_1, \\ 1+\tfrac{1}{2}(b_4-c_1-b_1), \tfrac{1}{2}(1+b_4-c_1-b_1); \\ 1+b_4-c_1-b_1, 1-\tfrac{1}{2}(c_1+b_1), \tfrac{1}{2}(3-c_1-b_1), \\ 2+c_2+c_3-2b_1-c_1 \end{bmatrix}$$
$$+\Gamma\begin{bmatrix} a_2, a_3, a_4, 1-a_2, 1-a_3, 1-a_4, 1+c_2+c_3-c_1-2b_1, \\ 3b_1+b_4+c_1-2c_2-2c_3-1; \\ b_1, c_1+b_1-c_2, c_1+b_1-c_3, b_4, 1-b_1, 1+c_2-c_1-b_1, \\ 1+c_3-c_1-b_1, b_1+b_4-c_2-c_3, 3b_1+c_1-2c_2-2c_3 \end{bmatrix}$$
$$(2.3) \quad \times\, {}_5F_4\begin{bmatrix} c_1+b_1-c_2-c_3, b_1-c_2, b_1-c_3, \\ \tfrac{1}{2}(3b_1+b_4+c_1-2c_2-2c_3-1), \tfrac{1}{2}(3b_1+b_4+c_1-2c_2-2c_3); \\ 2b_1+c_1-c_2-c_3, b_1+b_4-c_2-c_3, \\ \tfrac{1}{2}(3b_1+c_1-2c_2-2c_3), \tfrac{1}{2}(1+3b_1+c_1-2c_2-2c_3) \end{bmatrix}$$

$$+\,\Gamma\left[\begin{matrix}2-a_2, a_2-1, 1+a_3-a_2, 1+a_4-a_2, a_2-a_3, a_2-a_4;\\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, 1+c_1+b_1-c_3-a_2,\\ 1+b_4-a_2, a_2-c_1, a_2-c_2, a_2-c_3, a_2\end{matrix}\right]$$

$$\times\, {}_4H_4\left[\begin{matrix}1+c_1-a_2, 1+c_2-a_2, 1+c_3-a_2, 1-a_2;\\ 1+b_1-a_2, 1+c_1+b_1-c_2-a_2, 1+c_1+b_1-c_3-a_2,\\ 1+b_4-a_2\end{matrix}\right]$$

$$+ \text{ idem } (a_2; a_3, a_4) = 0.$$

If we put $b_4 = a_4$, $a_2 = c_1 + b_1 - c_2$ and $a_3 = c_1 + b_1 - c_3$ in (2.3), we get a relation between two Saalschützian ${}_5F_4$, two nearly-poised ${}_4F_3$ of the second kind and a nearly-poised ${}_4F_3$ series of the first kind. Also, if in this new relation we put $c_1 = 1 + c_3 - b_1 + n$, we get a relation between a terminating nearly-poised ${}_4F_3$ series of the second kind and a terminating Saalschützian ${}_5F_4$ series and, after reversing the terminating Saalschützian ${}_5F_4$ series, we get

$$(2.4)\qquad {}_4F_3\left[\begin{matrix}c_3-n, 1+c_3-b_1, c_2-n, -n;\\ b_1-n, 1+c_3-c_2, a_4-n\end{matrix}\right]$$

$$= \frac{(a_4-c_3)_n}{(a_4-n)_n}\, {}_5F_4\left[\begin{matrix}1+c_3-a_4, b_1-c_2, \tfrac{1}{2}(c_3-n), \tfrac{1}{2}(1+c_3-n), -n;\\ b_1-n, 1+c_3-c_2, \tfrac{1}{2}(1+c_3-a_4-n),\\ 1+\tfrac{1}{2}(c_3-a_4-n)\end{matrix}\right]$$

which is § 4.5 (1) of Bailey **(1)**.

Again, if we reverse the first ${}_4H_4$ series on the left of (2.3) and then put $a_2 = 1$, we get

$$ {}_4F_3\left[\begin{matrix}1-b_4, 1-b_1, 1+c_2-c_1-b_1, 1+c_3-c_1-b_1;\\ 1-c_1, \quad 1-c_2, \quad 1-c_3\end{matrix}\right]$$

$$= \Gamma\left[\begin{matrix}1-c_1, 1-c_2, 1-c_3, 2b_1+c_1-c_2-c_3-1;\\ c_1+b_1-c_2-c_3, b_1-c_2, b_1-c_3, 2-b_1-c_1\end{matrix}\right]$$

$$\times\, {}_5F_4\left[\begin{matrix}1-b_1, 1+c_2-c_1-b_1, 1+c_3-c_1-b_1, \tfrac{1}{2}(1+b_4-c_1-b_1),\\ 1+\tfrac{1}{2}(b_4-c_1-b_1);\\ 1+b_4-c_1-b_1, 1-\tfrac{1}{2}(c_1+b_1), \tfrac{1}{2}(3-c_1-b_1),\\ 2+c_2+c_3-2b_1-c_1\end{matrix}\right]$$

$$(2.5)\qquad +\,\Gamma\left[\begin{matrix}1-c_1, 1-c_2, 1-c_3, 1+c_2+c_3-c_1-2b_1,\\ 3b_1+b_4+c_1-2c_2-2c_3-1;\\ 1-b_1, 1+c_2-c_1-b_1, 1+c_3-c_1-b_1, b_1+b_4-c_2-c_3,\\ 3b_1+c_1-2c_2-2c_3\end{matrix}\right]$$

$$\times\, {}_5F_4\left[\begin{matrix}c_1+b_1-c_2-c_3, b_1-c_2, b_1-c_3, \tfrac{1}{2}(3b_1+b_4+c_1-2c_2-2c_3),\\ \tfrac{1}{2}(3b_1+b_4+c_1-2c_2-2c_3-1);\\ 2b_1+c_1-c_2-c_3, b_1+b_4-c_2-c_3, \tfrac{1}{2}(3b_1+c_1-2c_2-2c_3),\\ \tfrac{1}{2}(1+3b_1+c_1-2c_2-2c_3)\end{matrix}\right]$$

which is (1) of § 6.5 of **(1)**.

3. In this section, I consider certain transformations of nearly-poised series of the type ${}_5H_5$. If we take $M = 5$ in (2.1) and then reverse the first ${}_5H_5$ series in it and put $C_5 = 0$, we get the following relation between a nearly-poised ${}_5F_4$ series of the first kind and four nearly-poised ${}_5H_5$ series:

$$\Gamma\left[\begin{matrix} a_2, a_3, a_4, a_5, 1 - a_2, 1 - a_3, 1 - a_4, 1 - a_5; \\ b_1, c_1 + b_1 - c_2, c_1 + b_1 - c_3, c_1 + b_1 - c_4, b_5, 1 - c_1, 1 - c_2, \\ 1 - c_3, 1 - c_4 \end{matrix}\right]$$

$$\times\, {}_5F_4\left[\begin{matrix} 1 - b_5, 1 - b_1, 1 + c_2 - c_1 - b_1, 1 + c_3 - c_1 - b_1, \\ 1 + c_4 - c_1 - b_1; \\ 1 - c_1, 1 - c_2, 1 - c_3, 1 - c_4 \end{matrix}\right]$$

$$(3.1) \qquad + \Gamma\left[\begin{matrix} 2 - a_2, a_2 - 1, 1 + a_3 - a_2, 1 + a_4 - a_2, 1 + a_5 - a_2, a_2 - a_3, \\ a_2 - a_4, a_2 - a_5; \\ 1 + b_1 - a_2, 1 \div c_1 + b_1 - c_2 - a_2, 1 + c_1 + b_1 - c_3 - a_2, \\ 1 + c_1 + b_1 - c_4 - a_2, \\ 1 + b_5 - a_2, a_2 - c_1, a_2 - c_2, a_2 - c_3, a_2 - c_4, a_2 \end{matrix}\right]$$

$$\times\, {}_5H_5\left[\begin{matrix} 1 + c_1 - a_2, 1 + c_2 - a_2, 1 + c_3 - a_2, 1 + c_4 - a_2, 1 - a_2; \\ 1 + b_1 - a_2, 1 + c_1 + b_1 - c_2 - a_2, 1 + c_1 + b_1 - c_3 - a_2, \\ 1 + c_1 + b_1 - c_4 - a_2, 1 + b_5 - a_2 \end{matrix}\right]$$

$$+ \text{ idem } (a_2; a_3, a_4, a_5) = 0.$$

Now if we first put $a_2 = c_1 + b_1 - c_2$ in (3.1) and then in the new relation put $c_1 = 2c_2 - b_1 - 1$, we get the following relation between a nearly-poised ${}_5F_4$ series of the first kind and three nearly-poised ${}_5H_5$ series:

$$\Gamma\left[\begin{matrix} a_3, a_4, a_5, 2 - c_2, 1 - a_3, 1 - a_4, 1 - a_5; \\ b_1, 2c_2 - c_3 - 1, 2c_2 - c_4 - 1, b_5, 2 + b_1 - 2c_2, 1 - c_2, 1 - c_3, 1 - c_4 \end{matrix}\right]$$

$$\times\, {}_5F_4\left[\begin{matrix} 1 - b_5, 1 - b_1, 2 - c_2, 2 + c_3 - 2c_2, 2 + c_4 - 2c_2; \\ 2 + b_1 - 2c_2, 1 - c_2, 1 - c_3, 1 - c_4 \end{matrix}\right]$$

$$(3.2) \qquad + \Gamma\left[\begin{matrix} 2 - a_3, a_3 - 1, 1 + a_4 - a_3, 1 + a_5 - a_3, 1 + a_3 - c_2, \\ a_3 - a_4, a_3 - a_5; \\ 1 + b_1 - a_3, 2c_2 - c_3 - a_3, 2c_2 - c_4 - a_3, 1 + b_5 - a_3, 1 + b_1 \\ + a_3 - 2c_2, a_3 - c_2, a_3 - c_3, a_3 - c_4, a_3 \end{matrix}\right]$$

$$\times\, {}_5H_5\left[\begin{matrix} 2c_2 - b_1 - a_3, 1 + c_2 - a_3, 1 + c_3 - a_3, 1 + c_4 - a_3, 1 - a_3; \\ 1 + b_1 - a_3, c_2 - a_3, 2c_2 - c_3 - a_3, 2c_2 - c_4 - a_3, 1 + b_5 - a_3 \end{matrix}\right]$$

$$+ \text{ idem } (a_3; a_4, a_5) = 0.$$

Since the nearly-poised ${}_5F_4$ series of the first kind on the left of (3.2) can be expressed in terms of two Saalschützian ${}_5F_4$ series [cf. (**1**, § 6.5)], we get the following relation between two Saalschützian ${}_5F_4$ series and three nearly-poised ${}_5H_5$ series:

$$\Gamma\left[\begin{matrix} a_3, a_4, a_5, 2 - c_2, 1 - a_3, 1 - a_4, 1 - a_5, 2c_2 + b_1 - c_3 - c_4 - 2; \\ b_1, 2c_2 - c_3 - 1, 2c_2 - c_4 - 1, b_5, 1 - c_2, b_1 - c_3, b_1 - c_4, \\ \qquad 2c_2 - c_3 - c_4 - 1,\ 3 - 2c_2 \end{matrix}\right]$$

$$\times {}_5F_4\left[\begin{matrix} 1 - b_1, 2 + c_3 - 2c_2, 2 + c_4 - 2c_2, \tfrac{1}{2}(1 + b_5 - 2c_2), 1 + \tfrac{1}{2} \\ \qquad (b_5 - 2c_2); \\ 2 + b_5 - 2c_2, 1 - c_2, \tfrac{3}{2} - c_2, 3 + c_3 + c_4 - b_1 - 2c_2 \end{matrix}\right]$$

$$+ \frac{(1 + b_5 - 2c_2)}{2(1 - c_2)}\Gamma\left[\begin{matrix} a_3, a_4, a_5, 2 - c_2, 1 - a_3, 1 - a_4, 1 - a_5, 2c_2 + b_5 \\ \quad + 2b_1 - 2c_3 - 2c_4 - 3, 2 + c_3 + c_4 - b_1 - 2c_2; \\ b_1, 2c_2 - c_3 - 1, 2c_2 - c_4 - 1, b_5, 1 - b_1, 2 + c_3 \\ \quad - 2c_2, 2 + c_4 - 2c_2, 1 - c_2, b_5 + b_1 - c_3 - c_4, \\ \quad 2(c_2 + b_1 - c_3 - c_4 - 1) \end{matrix}\right] \tag{3.3}$$

$$\times {}_5F_4\left[\begin{matrix} b_1 - c_3, b_1 - c_4, 2c_2 - c_3 - c_4 - 1, c_2 + b_1 - c_3 - c_4 \\ \qquad + \tfrac{1}{2}(b_5 - 3), c_2 + b_1 + \tfrac{1}{2}b_5 - c_3 - c_4 - 1; \\ b_5 + b_1 - c_3 - c_4, c_2 + b_1 - c_3 - c_4 - 1, c_2 - c_3 - c_4 + b_1 - \tfrac{1}{2}, \\ 2c_2 + b_1 - c_3 - c_4 - 1 \end{matrix}\right]$$

$$+ \Gamma\left[\begin{matrix} 2 - a_3, a_3 - 1, 1 + a_4 - a_3, 1 + a_5 - a_3, 1 + a_3 - c_2, \\ \qquad a_3 - a_4, a_3 - a_5; \\ 1 + b_1 - a_3, 2c_2 - c_3 - a_3, 2c_2 - c_4 - a_3, 1 + b_5 - a_3, 1 + b_1 \\ \quad + a_3 - 2c_2, a_3 - c_2, a_3 - c_3, a_3 - c_4, a_3 \end{matrix}\right]$$

$$\times {}_5H_5\left[\begin{matrix} 2c_2 - b_1 - a_3, 1 + c_2 - a_3, 1 + c_3 - a_3, 1 + c_4 - a_3, 1 - a_3; \\ 1 + b_1 - a_3, c_2 - a_3, 2c_2 - c_3 - a_3, 2c_2 - c_4 - a_3, 1 + b_5 - a_3 \end{matrix}\right]$$

$$+ \text{idem } (a_3;\ a_4, a_5) = 0.$$

If we take $b_5 = a_5$, $a_4 = 2c_2 - c_4 - 1$ and $a_3 = 2c_2 - c_3 - 1$ in (3.3), we get a relation between two Saalschützian ${}_5F_4$, two nearly-poised ${}_5F_4$ of the second kind and a nearly-poised ${}_5F_4$ series of the first kind. Also, if we put $2 + c_3 - 2c_2 = -n$ in this new transformation, we get a relation between a terminating nearly-poised ${}_5F_4$ series of the second kind and a terminating Saalschützian ${}_5F_4$ series and after reversing the terminating Saalschützian ${}_5F_4$ series, we get

$$ {}_5F_4\left[\begin{matrix} c_3 - n, 1 + \tfrac{1}{2}(c_3 - n), 1 + c_3 - b_1, c_4 - n, -n; \\ \tfrac{1}{2}(c_3 - n), b_1 - n, 1 + c_3 - c_4, a_5 - n \end{matrix}\right] \tag{3.4}$$

$$= \frac{(a_5 - c_3 - 1 - n)(a_5 - c_3)_{n-1}}{(a_5 - n)_n}$$

$$\times {}_5F_4\left[\begin{matrix} b_1 - c_4, 1 + c_3 - a_5, 1 + \tfrac{1}{2}(c_3 - n), \tfrac{1}{2}(1 + c_3 - n), -n; \\ b_1 - n, 1 + c_3 - c_4, \tfrac{1}{2}(3 + c_3 - a_5 - n), 1 + \tfrac{1}{2}(c_3 - a_5 - n) \end{matrix}\right]$$

which is (2) of § 4.5 of **(1)**.

Again, if we reverse the first ${}_5H_5$ series in (3.3) and put $a_3 = 1$, we get the following relation:

$$(3.5)\qquad {}_5F_4\left[\begin{matrix}1-b_5, 2-c_2, 1-b_1, 2+c_3-2c_2, 2+c_4-2c_2;\\ 1-c_2, 2+b_1-2c_2, 1-c_3, 1-c_4\end{matrix}\right]$$

$$=\Gamma\left[\begin{matrix}2+b_1-2c_2, 1-c_3, 1-c_4, 2c_2+b_1-c_3-c_4-2;\\ b_1-c_3, b_1-c_4, 2c_2-c_3-c_4-1, 3-2c_2\end{matrix}\right]$$

$$\times {}_5F_4\left[\begin{matrix}1-b_1, 2+c_3-2c_2, 2+c_4-2c_2, \tfrac{1}{2}(1+b_5-2c_2), 1+\tfrac{1}{2}(b_5-2c_2);\\ 2+b_5-2c_2,\ 1-c_2,\ \tfrac{3}{2}-c_2,\ 3+c_3+c_4-b_1-2c_2\end{matrix}\right]$$

$$+\frac{(1+b_5-2c_2)}{2(1-c_2)}\,\Gamma\left[\begin{matrix}2c_2+b_5+2b_1-2c_3-2c_4-3,\\ 2+c_3+c_4-b_1-2c_2, 2+b_1-2c_2, 1-c_3, 1-c_4;\\ 1-b_1, 2+c_3-2c_2, 2+c_4-2c_2,\\ \qquad b_5+b_1-c_3-c_4, 2(c_2+b_1-c_3-c_4-1)\end{matrix}\right]$$

$$\times {}_5F_4\left[\begin{matrix}b_1-c_3, b_1-c_4, 2c_2-c_3-c_4-1, c_2+b_1+\tfrac{1}{2}(b_5-3)\\ \qquad -c_3-c_4, c_1+b_1+\tfrac{1}{2}b_5-c_3-c_4-1;\\ b_5+b_1-c_3-c_4,\ c_2+b_1-c_3-c_4-1,\ c_1+b_1-c_3-c_4-\tfrac{1}{2},\\ 2c_2+b_1-c_3-c_4-1\end{matrix}\right]$$

which is the generalization of (2) of §4.6 of **(1)**.

4. The sum of a nearly-poised ${}_3H_3$. Taking $M = 3$ and $a_2 = c_1 + b_1 - c_2$ in (2.1) and then putting $c_1 = 2c_2 - b_1 - 1$ in the new transformation, we get the following relation between two nearly-poised ${}_3H_3$ series:

$$(4.1)\qquad {}_3H_3\left[\begin{matrix}2c_2-b_1-1, c_2, c_3;\\ b_1, c_2-1, b_3\end{matrix}\right]$$

$$=\Gamma\left[\begin{matrix}1+a_3-c_2, b_1, b_3, 2+b_1-2c_2, 1-c_2, 1-c_3;\\ 1+b_1-a_3, 1+b_3-a_3, 1+b_1+a_3-2c_2, a_3-c_2, a_3-c_3, 2-c_2\end{matrix}\right]$$

$$\times {}_3H_3\left[\begin{matrix}2c_2-b_1-a_3, 1+c_2-a_3, 1+c_3-a_3;\\ 1+b_1-a_3,\quad c_2-a_3,\quad 1+b_3-a_3\end{matrix}\right].$$

If we put $b_3 = a_3$ in (4.1), we get a relation between a nearly-poised ${}_3H_3$ and a summable nearly-poised ${}_3F_2$ series [cf. § 6.4 (2) of **(1)**]. Hence, we get

$$(4.2)\qquad {}_3H_3\left[\begin{matrix}2c_2-b_1-1, c_2, c_3;\\ b_1,\ c_2-1,\ a_3\end{matrix}\right]$$

$$=\frac{(1+a_3+c_3-2c_2)}{2(1-c_2)}\,\Gamma\left[\begin{matrix}b_1, a_3, 2+b_1-2c_2, 1-c_3,\\ \qquad 2b_1+a_3-c_3-2c_2-1;\\ 1+b_1+a_3-2c_2, a_3-c_3,\\ \qquad b_1-c_3, 2(b_1-c_2)\end{matrix}\right].$$

If we put $b_1 = 1$ in (4.2), we get

$$(4.3)\qquad {}_3F_2\left[\begin{matrix}2(c_2-1), c_2, c_3;\\ c_2-1, a_3\end{matrix}\right]$$

$$= (1+a_3+c_3-2c_2)\,\Gamma\left[\begin{matrix}a_3, 1+a_3-c_3-2c_2;\\ 2+a_3-2c_2, a_3-c_3\end{matrix}\right].$$

Again, if we put $c_3 = -n$ in (4.3), we get

$$(4.4)\qquad {}_3F_2\left[\begin{matrix}2(c_2-1), c_2, -n;\\ c_2-1, a_3\end{matrix}\right] = \frac{(1+a_3-2c_2-n)(2+a_3-2c_2)_{n-1}}{(a_3)_n}$$

which is § 4.5 (1.1) of **(1)**.

Also, if we put $a_3 = 1$ in (4.2), we again get the sum of a nearly-poised ${}_3F_2$ series of the first kind.

It may be remarked that (4.2) can be obtained more easily by using the identity*

$$(4.5)\qquad K \times {}_3H_3\left[\begin{matrix}1+\frac{1}{2}K, b, a;\\ \frac{1}{2}K, 1+K-b, w\end{matrix}\right]$$
$$= 2(K-b)\,{}_2H_2\left[\begin{matrix}b, a;\\ K-b, w\end{matrix}\right] - (K-2b)\,{}_2H_2\left[\begin{matrix}b, a;\\ 1+K-b, w\end{matrix}\right]$$

and summing the two ${}_2H_2$ series on the right of (4.5).

I am grateful to Dr. R. P. Agarwal for his kind guidance during the preparation of this paper.

*I am grateful to the referee for pointing this out.

References

1. W. N. Bailey, *Generalised Hypergeometric Series* (Cambridge 1935).
2. L. J. Slater, Quart. J. Math. (2) *3* (1952), 73–80.
3. H. S. Shukla, *Certain Transformations of nearly-poised bilateral hypergeometric series,* Ganita, *7* (1956).
4. ———, *Certain Transformations of nearly-poised basic bilateral hypergeometric series of the type* ${}_M\Psi_M$, to appear in Math. Z.

Lucknow University

COMMA-FREE CODES

S. W. GOLOMB, BASIL GORDON AND L. R. WELCH

1. A General Combinatorial Problem. Let n be a fixed positive integer, and consider an alphabet consisting of the numbers $1, 2, \ldots, n$. With this alphabet form all possible k-letter words $(a_1\, a_2 \ldots a_k)$, where k is also fixed. There are evidently n^k such words in all.

Definition: A set D of k-letter words is called a *comma-free dictionary* if whenever $(a_1\, a_2 \ldots a_k)$ and $(b_1\, b_2 \ldots b_k)$ are in D, the "overlaps" $(a_2\, a_3 \ldots a_k\, b_1)$, $(a_3 \ldots a_k\, b_1\, b_2), \ldots, (a_k\, b_1 \ldots b_{k-1})$ are not in D.

The problem to be investigated here is that of determining the greatest number of words that a comma-free dictionary can possess. We denote this number by $W_k(n)$.

THEOREM 1.

$$W_k(n) < \frac{1}{k} \sum \mu(d)\, n^{k/d},$$

where the summation is extended over all divisors d of k, and $\mu(d)$ is the Möbius function, defined by

$$\mu(d) = \begin{cases} 1 & \textit{if } d = 1 \\ 0 & \textit{if } d \textit{ has any square factor} \\ (-1)^r & \textit{if } d = p_1\, p_2 \ldots p_r, \textit{ where } p_1, \ldots, p_r \textit{ are distinct primes.} \end{cases}$$

Proof. Let d be a divisor of k. We say that a word $(a_1\, a_2 \ldots a_k)$ has subperiod d if it is of the form $(a_1\, a_2 \ldots a_d\, a_1\, a_2 \ldots a_d \ldots a_1\, a_2 \ldots a_d)$, and if d is the smallest number for which this is true. For example, if $k = 6$, then $(a\, a\, a\, a\, a\, a)$ has subperiod 1, $(a\, b\, a\, b\, a\, b)$ has subperiod 2 if $a \neq b$, $(a\, b\, c\, a\, b\, c)$ has subperiod 3 if $a \neq b$ or $b \neq c$, and all other words have subperiod 6. Any word w of a comma-free dictionary must have subperiod k because otherwise ww would contain an overlap of w. (Consider for example, $[a\, b\, c\, a\, b\, c]\, [a\, b\, c\, a\, b\, c]$.) We shall call words of subperiod k *primitive*.

For later purposes it is convenient to call two words *equivalent* if one is a cyclic permutation of the other, and to speak of $(a_1\, a_2 \ldots a_k)$, $(a_2 \ldots a_k\, a_1), \ldots$, $(a_k\, a_1 \ldots a_{k-1})$ as forming an equivalence class. If $(a_1\, a_2 \ldots a_k)$ is primitive, then its equivalence class is also called primitive, and consists of k distinct words. At most one of these can be a word in D, for otherwise a contradiction

Received July 11, 1957.

would again arise upon considering the overlaps of ww. Hence, if $P_k(n)$ is the total number of primitive words, then

$$W_k(n) \leqslant \frac{1}{k} P_k(n).$$

But since each of the n^k words has some subperiod, $P_k(n)$ satisfies the equation

$$\sum_{d/k} P_d(n) = n^k,$$

from which we obtain

$$P_k(n) = \sum_{d/k} \mu(d)\, n^{k/d}$$

by Möbius inversion. For the Möbius inversion formula, cf. (**3**, p. 28).

2. Results for k odd. Theorem 1 gives a general upper bound for $W_k(n)$, of which a few examples are

$$W_1(n) \leqslant n, \quad \mathrm{W}_2(n) \leqslant \tfrac{1}{2}(n^2 - n), \quad W_3(n) \leqslant \tfrac{1}{3}(n^3 - n), \quad W_4(n) \leqslant \tfrac{1}{4}(n^4 - n^2).$$

In many cases this upper bound is actually attained. We believe this to be true for all odd k, and have proved it for all odd $k \leqslant 15$. Note, from the proof of Theorem 1, that the upper bound will be attained if and only if a word can be chosen from each primitive equivalence class so as to form a comma-free dictionary.

THEOREM 2. *For arbitrary n,*

$$W_k(n) = \frac{1}{k} \sum_{d/k} \mu(d)\, n^{k/d}$$

if $k = 1, 3, 5, 7, 9, 11, 13, 15.$

Proof. For $k = 1$, the proof that $W_1(n) = n$ is immediate. For the other values of k we shall show how to select a word from each primitive equivalence class in such a way that a comma-free dictionary is obtained.

(i) In the case $k = 3$, let D be the set of all words $(a\ b\ c)$ satisfying the inequalities $a < b > c$. It is immediately seen that D is comma free. In order to show that the number of words in D is $\frac{1}{3}(n^3 - n)$, one could, of course, count the number of solutions of the inequalities $a < b > c$, where a, b, c are integers between 1 and n. But it is simpler to observe that if $(a_1\, a_2\, a_3)$ is any primitive word (that is, one for which $a_1 = a_2 = a_3$ does not hold), then some cyclic permutation of it clearly satisfies $a < b > c$. In particular $W_3(4) = 20$, a fact which will be useful in section 5.

(ii) For $k = 5$, the procedure is similar but more complex. Let D consist of all words $(a\ b\ c\ d\ e)$ satisfying $a < b > c$, $d > e$, and also of all words satisfying $a < b < c < d > e$. It can be readily verified that D is comma-free. In order to show that the number of elements in D is the upper bound (in this case $\frac{1}{5}(n^5 - n)$), we must prove that every primitive equivalence class contains a word of D. For this purpose let $+$ denote any number which is > 0, and $-$ any number which is < 0. Using this notation the elements of D can be characterized as those words $(a\ b\ c\ d\ e)$ for which the sequence of differences $b - a$, $c - b$, $d - c$, $e - d$ is of one of the forms $+ - - -$, $+ - + -$, or $+ + + -$. These patterns are precisely those which begin with an odd number of $+$'s and end with an odd number of $-$'s, a property which we shall call property P. (Incidentally, in the case $k = 3$ our dictionary consisted of words $[a\ b\ c]$ for which the differences $b - a$, $c - b$ were of the form $+ -$, that is, possessed property P.) Given any primitive word $(p\ q\ r\ s\ t)$ we form the differences $q - p$, $r - q$, $s - r$, $t - s$, $p - t$ obtained by representing p, q, r, s, t as points on a circle. We call $p - t$ the improper difference. By performing a suitable cyclic permutation on $(p\ q\ r\ s\ t)$ we can arrange matters so that any one of the five differences becomes the improper one.

Now by primitivity, both $+$'s and $-$'s appear among the differences, and since the total number of signs is 5, there must occur someplace a run of $-$'s followed by a run of $+$'s, the lengths of these runs being of opposite parity (note that this result depends only on the fact that the total number of signs is *odd*). Permuting cyclically we can put the run of $+$'s at the beginning, the run of $-$'s at the end, and make the improper difference have the sign which occurred an even number of times. The proper differences will then satisfy property P, and hence, given any primitive word, some cyclic permutation of it is in D.

(iii) For $k = 7$ we use the same method. Every primitive word has some cyclic permutation with property P. Its proper differences will then have one of the following 8 patterns:

$$(+ + + + + -)\ (+ + + - + -)\ (+ + + - - -)\ (+ - + + + -)$$
$$(+ - + - - -)\ (+ - - + + -)\ (+ - - - + -)\ (+ - - - - -)$$

Letting D consist of all such words, we find that D is comma-free. (The verification begins to become tedious, but is straightforward. The first overlap of two words in D begins with an even number of $+$'s, hence is not in D, the second overlap ends with a $+$, etc.)

(iv) When $k = 9$, the difficulty arises that there may be more than one word in a primitive equivalence class with property P. This happens for words $(a_1\ a_2\ a_3\ a_4\ a_5\ a_6\ a_7\ a_8\ a_9)$ with $a_1 < a_2 > a_3$, $a_4 < a_5 > a_6$, $a_7 < a_8 > a_9$. Here the permutations $(a_4\ a_5\ a_6\ a_7\ a_8\ a_9\ a_1\ a_2\ a_3)$ and $(a_7\ a_8\ a_9\ a_1\ a_2\ a_3\ a_4\ a_5\ a_6)$ also have property P. But notice that these words consist of three blocks of three letters, each of the type used for $k = 3$. This suggests the idea of ordering the 3-letter words $(a\ b\ c)$ with $a < b > c$ in some fashion (say lexicographically), and choosing for the dictionary D that one of the three possibilities which is of the form $w_1 < w_2 > w_3$ in this ordering. For example, in the case of the word $(1\ 3\ 1\ 1\ 2\ 2\ 2\ 3\ 1)$, the permutation $(1\ 2\ 2\ 2\ 3\ 1\ 1\ 3\ 1)$ would be selected for D, because $(122) < (231) > (131)$ if lexicographic ordering is employed. Adopting this convention, the dictionary which results is comma-free.

(v) For $k = 11$, 13, and 15 the same methods can be used, but the work becomes increasingly cumbersome. It is conceivable that all odd k can be treated in this manner, but we have stopped with the proof that $15W_{15}(n) = n^{15} - n^5 - n^3 + n$.

This case, the first where k has two distinct prime factors, is particularly powerful evidence for the validity of the general conjecture.

3. Results for Even k.

When k is even, the results are much less complete, and we cannot even formulate a plausible conjecture as to the value of $W_k(n)$. We begin with

THEOREM 3. *$W_2(n) = [\frac{1}{3}n^2]$, where $[x]$ denotes the integral part of x.*

Proof. Let D be any comma-free dictionary, and define A to be the set of all integers which begin some word of D but never end a word of D. Similarly, let B be the set of integers which both begin and end words of D, and C the set of integers which only end words of D. For example, if $D = \{(43), (41), (35), (25), (15)\}$, then

$$A = \{4, 2\}, \quad B = \{3, 1\}, \quad C = \{5\}.$$

D must evidently consist of words of the forms $(a\ b)$, $(a\ c)$, $(b_1\ b_2)$, or $(b\ c)$, where $a \in A$, $b, b_1, b_2 \in B$, and $c \in C$. But $(b_1 b_2)$ cannot occur, for there is some word in D ending in b_1, and some word beginning with b_2, and the comma-free property therefore excludes $(b_1\ b_2)$. This leaves only words of the forms $(a\ b)$, $(a\ c)$, $(b\ c)$, and it is immediately seen that the set of all these is comma-free. If α is the number of elements of A, β of B, and γ of C, then the number of words in D is at most $\alpha\beta + \beta\gamma + \gamma\alpha$. Maximizing the quantity $\alpha\beta + \beta\gamma + \gamma\alpha$ subject to the constraint $\alpha + \beta + \gamma = n$, we see that α, β, γ should be chosen as nearly equal as possible, in which case

$$\alpha\beta + \beta\gamma + \gamma\alpha = [\tfrac{1}{3}n^2].$$

For example, if $n = 3$ we would take $A = \{1\}$, $B = \{2\}$, $C = \{3\}$, and obtain $D = \{(12), (13), (23)\}$. It is not difficult to see that for arbitrary n we may choose D to be the set of all words w congruent to one of these words (mod 3), where

$$(a_1\ a_2 \ldots a_k) \equiv (b_1\ b_2 \ldots b_k) \pmod{m}$$

means

$$a_j \equiv b_j \pmod{m}, \qquad j = 1, 2, \ldots, k.$$

Thus for $n = 5$, $D = \{(12), (15), (42), (45), (13), (43), (23), (53)\}$.

THEOREM 4. *If k is any even integer, then the upper bound given by Theorem* 1 *is not attained by $W_k(n)$ provided that $n > 3^{\frac{1}{2}k}$.*

Proof. Let $k = 2j$, and let L be a comma-free dictionary. We define S_1 to be the set of all j-tuples $(a_1 a_2 \ldots a_j)$ which form the first half of some word in L, and S_2 to be the set of k-tuples $(a_{j+1}\, a_{j+2} \ldots a_k)$ which form the second half of some word in L. Then we put

$$A = S_1 \cap S_2', \quad B = S_1 \cap S_2, \quad C = S_1' \cap S_2, \quad D = S_1' \cap S_2',$$

where the prime denotes complementation. The four sets A, B, C, D are mutually exclusive and mutually exhaustive, so that any j-tuple is in one and only one of them. Hence to every k-letter word we may associate a pair (AA), (AB), ... or (DD) depending on which set its first half falls into and which set its second half falls into. As in the proof of Theorem 3, it is seen that for words in L the type (BB) cannot arise, and hence only (AB), (AC), and (BC) remain.

The upper bound of Theorem 1 was the number of primitive equivalence classes. To prove Theorem 4, we will show the existence of a primitive k-letter word, such that no cyclic permutation of it has any of the forms (AB), (AC), or (BC).

Consider the following particular blocks of length j:

$$(1, 1, 1, \ldots, 1, m) \qquad 1 \leqslant m \leqslant n.$$

Let T_i be the cyclic permutation which shifts each letter i units to the left. Define

$$F_m(i) = \begin{cases} 1 \textit{ if } T_i\,(1, 1, \ldots, m) \in A \cup D \\ 2 \textit{ if } T_i\,(1, 1, \ldots, m) \in B \\ 3 \textit{ if } T_i\,(1, 1, \ldots, m) \in C \end{cases}$$

For each m, $F_m(i)$ is a function with a domain of j elements and a range of 3 elements. There can be at most 3^j such functions, and since $n > 3^{\frac{1}{2}k} = 3^j$, there exist two distinct integers p and m such that $F_p \equiv F_m$ for all i. We now claim that no cyclic permutation of the word

$$w = (1\,1 \ldots p\,1\,1 \ldots m)$$

is of the form (AB), (AC), or (BC). For any permutation of w consists of a cyclic permutation of (1 1 ... p) followed by the *same* cyclic permutation of $(1\ 1 \ldots m)$ or vice versa. Since $F_p \equiv F_m$, we therefore get only the forms (AA), (AD), (DA), (DD), (BB), or (CC).

In particular, when $k = 4$, Theorem 4 proves that $4W_4(n) < n^4 - n^2$ for $n > 9$. By more delicate arguments it can be shown that this inequality is true for $n \geqslant 5$. On the other hand, if $n = 1, 2, 3$, then $4\ W_4(n) = n^4 - n^2$, as is seen by considering the dictionary D of words $(a\ b\ c\ d)$ satisfying $a < c$, $b \geqslant d$. The question of whether or not $W_4(4) = 60$ is still open. The best that can currently be proved is $W_4(4) \geqslant 56$.

4. Asymptotic Results. In this section we shall prove some theorems about the asymptotic behavior of $W_k(n)$ when k is fixed and $n \to \infty$.

THEOREM 5. *The limit*

$$\lim_{n\to\infty} \frac{W_k(n)}{n^k} = \alpha_k$$

exists.

Proof. We shall show that if n_0 is any fixed integer, then

$$\liminf_{n\to\infty} \frac{W_k(n)}{n^k} > \frac{W_k(n_0)}{n_0^k}.$$

This fact, coupled with the obvious boundedness of the ratio in question, proves the existence of the limit.

Consider then the integer n_0, and let D be a comma-free dictionary containing $W_k(n_0)$ words. For any arbitrary n, form the set S of all words w such that

$$w \equiv w_0 \pmod{n_0},$$

where $w_0 \in D$. (The definition of congruence is given after Theorem 3 together with an example of the present procedure.) S is clearly comma-free, and so if it contains $S_k(n)$ elements, then

$$W_k(n) > S_k(n).$$

But it is easy to see that

$$\lim_{n\to\infty} \frac{S_k(n)}{n^k} = \frac{W_k(n_0)}{n_0^k}.$$

This completes the proof of the theorem.

THEOREM 6. *If k is odd, then $\alpha_k = 1/k$.*

Proof. By Theorem 1,

$$W_k(n) < \frac{1}{k} \sum_{d/k} \mu(d)\, n^{k/d}.$$

If k is fixed and $n \to \infty$, the right hand side is asymptotically n^k/k. Hence,

$$\lim_{n\to\infty} \frac{W_k(n)}{n^k} < \frac{1}{k}.$$

On the other hand, consider the dictionary D defined as follows: Put $k = 2j - 1$ and let D consist of all words $(a_1\, a_2\, \ldots\, a_k)$ such that a_j is greater than any of the other a_i's. D is comma-free, as is easily verified, and the number of elements in D is equal to

$$\sum_{m=1}^{n-1} m^{k-1} \sim \frac{n^k}{k}$$

This shows that

$$\lim_{n\to\infty} \frac{W_k(n)}{n^k} > \frac{1}{k},$$

and thus establishes Theorem 6.

THEOREM 7. *If k is even, then $1/ek < \alpha_k < 1/k$.*

Proof. The first part of Theorem 6 holds for any fixed k. Hence, $\alpha_k < 1/k$. To obtain a lower bound, we divide the integers from 1 to n into two disjoint classes U and V. Then let D be the set of all words $(a_1 a_2 \ldots a_k)$ such that $a_1 \in U$ and $a_2, \ldots, a_k \in V$. D is clearly comma-free, and if the number of elements in V is v, then D contains $(n - v)v^{k-1}$ words. If v could take on all real values, then the maximum of this expression would occur for

$$v = \frac{k-1}{k} n ,$$

and would have the value

$$\frac{n^k}{k}\left(1 - \frac{1}{k}\right)^{k-1} .$$

The fact that v must be an integer has no effect, since taking

$$v = \left[\frac{k-1}{k} n\right]$$

gives a lower bound for $W_k(n)$ which is still asymptotically

$$\frac{n^k}{k}\left(1 - \frac{1}{k}\right)^{k-1} .$$

Hence

$$\alpha_k > \frac{1}{k}\left(1 - \frac{1}{k}\right)^{k-1} > \frac{1}{ek} .$$

For $k = 4$, Theorem 7 gives the bounds

$$\frac{27}{256} < \alpha_4 < \frac{1}{4} .$$

A better bound can be obtained from Theorem 5. As shown after Theorem 4, $W_4(3) = 18$, and hence

$$\alpha_4 > \frac{W_4(3)}{3^4} = \frac{18}{81} = \frac{2}{9} .$$

The exact value of α_k for even k is still an open question.

5. Applications. From their researches in the transfer of genetic information from parent to offspring, Crick, Griffith, and Orgel **(1)** advance the following hypothesis. Genetic information, they suggest, is encoded into a giant molecule (chromosome) by means of an affixed sequence of nucleotides, of which there are four types. Each such sequence is uniquely decodeable into a new protein molecule, consisting of a long sequence of amino acids, of which there are twenty types. They propose that each amino acid is specified by three consecutive nucleotides. However, only twenty of the sixty-four sequences of three nucleotides "make sense." Crick, Griffith, and Orgel theorize that the twenty sequences of nucleotides actually corresponding to amino acids form a comma-free dictionary. As we have seen, $W_3(4) = 20$, which agrees with the number of amino acids. The reasonableness of this condition can be seen if we think of the sequence of nucleotides as an infinite message, written without punctuation, from which any finite portion must be decodeable into a sequence of amino acids by suitable insertion of commas. If the manner of inserting commas were not unique, genetic chaos could result.

In their search for optimum coding techniques, Shannon, McMillan, and others have studied codes which are uniquely decipherable *in the large*—that is, when the entire message is available. This is a larger class than the comma-free messages, which must be uniquely decipherable *in the small*. In communications applications where only disjointed portions of a message are likely to be received, comma-free codes may indeed be useful. An excellent discussion of codes uniquely decipherable in the large is presented in (2).

References

1. H. C. Crick, J. S. Griffith, and L. E. Orgel, *Codes Without Commas*, Proc. Nat. Acad. Sci., *43* (1957), 416–421.
2. B. McMillan, *Two Inequalities Implied by Unique Decipherability*, IRE Transactions on Information Theory, *2* (1956), 115–116.
3. T. Nagell, *Introduction to Number Theory* (Uppsala, 1951).

ON THE NUMBER OF ORDINARY LINES DETERMINED BY n POINTS

L. M. KELLY AND W. O. J. MOSER

1. Introduction. More than sixty years ago, Sylvester **(13)** proposed the following problem: Let n given points have the property that the straight line joining any two of them passes through a third point of the set. Must the n points all lie on one line?

An alleged solution (not by Sylvester) advanced at the time proved to be fallacious and the problem remained unsolved until about 1933 when it was revived by Erdős **(7)** and others. Gallai (see **5**), Robinson (see **12**), Steinberg (see **4**, p. 30), Kelly (see **3**) and Lang **(11)** produced solutions of varying characters, the first affine, the second likewise affine (after dualizing), the third projective, and the fourth and fifth Euclidean. The answer is that in real projective space the points must indeed be on a line. Simple examples show that such is not the case in the complex projective plane **(3)**. The answer is also negative in finite projective geometries, where each line contains the same number of points. The property is very strongly dependent on the axioms of order.

The problem may be formulated in more general terms. Let P be a set of n points in real projective space and S the set of *connecting* lines which join these points. Call a line of S *ordinary* if it contains exactly two points of P. If S contains more than one line, show that it contains at least one ordinary line and determine lower bounds for the number of such lines. It is clear that the number of ordinary lines is invariant under a suitable central projection and so the question need only be settled in the real projective plane. The remainder of this investigation will be in the real projective plane, with P a set of n non-collinear points and S the set of their connecting lines. Let m denote the number of lines which are ordinary. Dirac **(6)** showed that $m > 3$ and Motzkin **(12)** showed that the order of magnitude of m is at least $\sqrt{n}$. It is the purpose of this note to show that $m > 3n/7$ and that, in a certain sense, this is a best possible bound.

2. Definitions, notation and preliminary theorems. A generic point of P is denoted by p and a generic line of S by s. Subscripts distinguish particular points and lines.

It is a known and easily established fact that a set of two or more lines in the plane which do not form a pencil effect a subdivision of the plane into two or more regions (see **(14)** for relevant definitions). With this in mind it is

Received May 21, 1957.

apparent that, except in the cases to be noted presently, the lines of S not passing through p dissect the plane into polygonal regions. In the event that the $n - 1$ points of P distinct from p are on a line, no division is effected. If exactly $n - 1$ of the points, including p, are on a line, then the division is into $n - 2$ angular regions, that is, regions bounded by two lines. In all other cases the division is into polygonal regions (bounded by at least three edges).

The point p is, of course, in the interior of one of these regions, which is called the ***residence*** of p, and p is said to ***reside*** in the region. The lines of S containing the edges of the residence are ***neighbours*** of p.

A set of n lines in the plane exactly $n - 1$ of which are concurrent is a ***near-pencil.*** This configuration is slightly exceptional in this study. We observe that if $n - 1$ points of P lie on one line, then S is a near-pencil.

THEOREM 2.1. *If a point p has precisely one neighbour, then S is a near-pencil.*

Proof. In this case the neighbour of p is the only line of S which does not pass through p; on this neighbour lie the remaining $n - 1$ points of P.

THEOREM 2.2. *If a point p has precisely two neighbours, then S is a near-pencil.*

Proof. In this case the lines of S which do not pass through p form a pencil; for otherwise they would form a proper dissection of the plane and p would have at least three neighbours. Let q be the vertex of this pencil. Let s_i, s_j be any two of the lines through q; let p_i, p_j, both different from q, be points on s_i, s_j respectively. The connecting line through p_i and p_j does not pass through q; hence, it must pass through p. It follows that there is only one line of S which passes through p; the remaining lines pass through q, which is necessarily a point of P.

THEOREM 2.3. *If S is not a near-pencil then each point of P has at least three neighbours.*

Proof. If a point of P has only one or two neighbours, then, by Theorems 2.1 and 2.2, S is a near-pencil.

3. Ordinary lines. The number of ordinary lines passing through p is the ***order*** of p. The number of neighbours of p which are ordinary lines is the ***rank*** of p. The order plus the rank is the ***index.***

THEOREM 3.1. *If the order of p is zero then every neighbour of p is an ordinary line.*

Proof. Suppose, to the contrary, that the neighbour s of p passes through three points of P, say p_1, p_2, p_3. Let x be a point on s which lies on the boundary of the residence of p. Suppose the notation so chosen that $p_1x//p_2p_3$, that is, p_1 and x separate p_2 and p_3. Since p is of order zero, the connecting line through p and p_1 passes through a third point of P, say p_4. The lines p_2p_4 and p_3p_4

intersect both the segments determined by p and x on the line through them. Hence, x cannot be a point of p's residence, and we have a contradiction which proves the theorem.

It is now apparent that the residence of a point of order zero is one of the polygons into which the plane is dissected by the m ordinary lines. Furthermore, it is clear that a polygon in this dissection cannot contain in its interior two points of order zero. Since the m ordinary lines pass through at most $2m$ points of P and dissect the plane into at most $\binom{m}{2} + 1$ polygons, it follows that

$$\binom{m}{2} + 1 + 2m \geqslant n.$$

This is Motzkin's proof of

THEOREM 3.2.

$$\binom{m+2}{2} \geqslant n.$$

Note that

$$\binom{m+2}{2} < \tfrac{1}{2}(m+2)^2$$

so that the theorem shows that $m > \sqrt{[2n]} - 2$.

THEOREM 3.3. *The index of each point of P which is not of order two is at least three.*

Proof. First observe that the theorem is true when S is a near-pencil and dismiss this case from further consideration.

Case 1. The order of p is zero. Since S is not a near-pencil, P has at least three neighbours; by Theorem 3.1 they are all ordinary lines.

Case 2. The order of p is one. Let p_1 be the second point on the ordinary line through p. The proof of Theorem 3.1 shows that if a neighbour of p is not ordinary, then it passes through p_1. Since three neighbours of p cannot have a common point, it follows that if p has more than three neighbours then at least two of them are ordinary. On the other hand, if p has precisely three neighbours then two of them must be ordinary. For, in this case, if s_1 and s_2 are two non-ordinary neighbours of p then both pass through p_1, which is therefore a vertex of the triangular residence of p. If x, a boundary point of the residence of p, is on s_1, and p_1, p_2, p_3 three points on s_1 with the notation so chosen that $p_1p_2//xp_3$ then (as in the proof of Theorem 3.1) pp_3 is a second ordinary line through p. This contradiction shows that if p has precisely three neighbours, then at most one of them is non-ordinary.

Case 3. The order of p is at least three. Then the index of p is at least three.

THEOREM 3.4. *If a line s of S is a neighbour of three points p_1, p_2, p_3, then the points of P which lie on s are on the connecting lines determined by p_1, p_2, p_3.*

Proof. Clearly, three points which have a common neighbour cannot be collinear. Let the points of intersection of s with the line p_ip_j be x_k (i, j, k a permutation of 1, 2, 3). Suppose p, a point different from x_1, x_2, x_3, lies on s in the segment x_ix_j/x_k, that is, $x_ix_j//px_k$. Then, because of the lines pp_i and pp_j, s cannot be a neighbour of p_k. Thus, p must coincide with one of the points x_1, x_2, x_3.

COROLLARY 3.4. *A line of S is a neighbour of at most four points.*

Remark. It is easy to show that if s is a neighbour of exactly four points of P, then s joins two diagonal points of the complete quadrangle determined by the four points. Furthermore, s is then ordinary.

THEOREM 3.5. *If I_i is the index of point p_i, then*

$$m \geqslant \frac{1}{6} \sum_{i=1}^{n} I_i.$$

Proof. We count the number of ordinary lines by observing the index of each point of P. In this counting, a particular ordinary line may be counted at most six times, four times as a neighbour (Corollary 3.4) and twice because it passes through a point.

THEOREM 3.6. $m \geqslant 3n/7$.

Proof. Let k be the number of points of order two. Clearly

$$m \geqslant k.$$

By Theorems 3.3 and 3.5

$$m \geqslant \frac{3(n-k)+2k}{6}.$$

Eliminating k from these inequalities, we obtain the desired result.

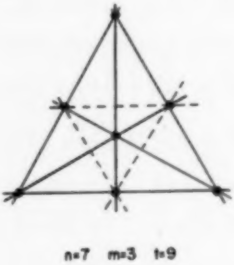

FIGURE 3.1

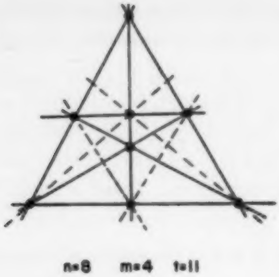

FIGURE 3.2

For $n = 7, 8$ the theorem shows that $m > 3, 4$. Figures 3.1 and 3.2 exhibit configuration of 7, 8 points with $m = 3, 4$ respectively (the ordinary lines are "broken"). In this sense Theorem 3.6 is best possible. However, for large n it would seem to us that the configuration with fewest possible lines is probably near the near-pencil arrangement. If this be so, then, for large n, m should be at least $n - 1$. Thus, a reasonable conjecture **(6)** is $m > \frac{1}{2}n$ for $n > 7$; the present method does not seem to allow us to draw this conclusion.

4. Connecting lines. In this section we derive an interesting inequality which we use to establish a bound on the number of connecting lines.

The dual of P is a set of n lines, $\bar{P}$, not a pencil; the dual of S is $\bar{S}$, the set of points of intersection of these lines.

In the dissection of the plane by the lines of $\bar{P}$, F_i denotes the number of polygons each having exactly i edges, and V_i denotes the number of vertices each incident with exactly i edges. V, E, and F denote the total number of vertices, edges and faces respectively.

Clearly $V_i = 0$ for all odd i,

4.10 $$V = V_4 + V_6 + V_8 + \dots$$

and

4.11 $$F = F_3 + F_4 + F_5 + \dots.$$

Since each edge has two vertices and belongs to two polygons

4.20 $$E = 2V_4 + 3V_6 + 4V_8 + \dots$$

and

4.21 $$2E = 3F_3 + 4F_4 + 5F_5 + \dots.$$

Adding 4.20 and 4.21 yields

4.22 $$3E = 2V_4 + 3V_6 + 4V_8 + \dots + 3F_3 + 4F_4 + 5F_5 + \dots.$$

By Euler's theorem,

4.3 $$V - E + F = 1.$$

Replacing V, F and E in 4.3 by their values 4.10, 4.11 and 4.22 respectively yields, after simplification,

4.4 $$V_4 = 3 + V_8 + 2V_{10} + 3V_{12} + \ldots + F_4 + 2F_5 + 3F_6 + \ldots .$$

Call a line of S which passes through precisely i points of P an *i-line* and let t_i denote the number of i-lines, for example, an ordinary line is a 2-line and $m = t_2$. Clearly, the dual of V_{2i} is t_i; hence, by dualizing 4.4 we establish

4.5 $$m = t_2 \geq 3 + t_4 + 2t_5 + 3t_6 + \ldots .$$

We have immediately Dirac's result **(6)**, $m \geq 3$.

Inequality 4.5 may be used to prove that, if n is even, $m > (n + 11)/6$. In fact, the assumption that there is a set P of n points with n even and $m \leq (n + 11)/6$ leads us to a contradiction. For, in such a case, we see from 4.5 that the number of points of P each of which is incident with a k-line for some $k \neq 3$ is at most

$$\begin{aligned}2t_2 + 4t_4 + 5t_5 + \ldots &\leq 2t_2 + 4(t_4 + 2t_5 + 3t_6 + \ldots) \\ &\leq 2t_2 + 4(t_2 - 3) = 6m - 12 \leq n - 1.\end{aligned}$$

Thus, there is at least one point of P incident solely with 3-lines; clearly n must be odd, and we have a contradiction

Call a point of P which is incident with exactly k connecting lines a k-point and let v_k denote the number of k-points. Clearly

4.60 $$\sum_{k=2} v_k = n;$$

also

4.61 $$\sum_{k=2} kt_k = \sum_{k=2} kv_k,$$

for in both sums a k-line is counted k times. From 4.5 we have

$$3t_2 + 3t_3 + 3t_4 + \ldots \geq 3 + 2t_2 + 3t_3 + 4t_4 + \ldots$$

which together with 4.61 yields

4.62 $$3t \geq 3 + \sum_{k=2} kt_k = 3 + \sum_{k=2} kv_k,$$

where

$$t = \sum t_k$$

denotes the number of connecting lines.

At the same time that Erdös **(7)** reproposed Sylvester's problem he also posed the following one: Show that S contains at least n lines. This was answered successfully by Steinberg (see **7**) as well as de Bruijn and Erdös **(5)** and Hanani **(9; 10)**. The configuration with S a near-pencil shows that this is a best possible result.

Erdös **(8)** conjectured that if n is large enough and at most $n - 2$ points are collinear then $2n - 4$ lines are determined. If we insist that at most $n - 3$

points be collinear, then approximately $3n$ lines should be determined. In general, if at most $n - k$ points are collinear and n is large (with respect to k) we would expect the minimum number of lines to be of order kn. This is the substance of the following theorem, a corollary of which establishes the truth of his conjecture.

THEOREM 4.1. *If at most $n - k$ points of P are collinear cnd*

4.70 $$n > \tfrac{1}{2}\{3(3k - 2)^2 + 3k - 1\}$$

then

$$t > kn - \tfrac{1}{2}(3k + 2)(k - 1).$$

The proof will follow that of

LEMMA 4.1. *If exactly $n - r$ points of P are on a line, and*

$$n > \frac{3r}{2} > 3$$

then

$$t > rn - \tfrac{1}{2}(3r + 2)(r - 1).$$

Proof. Suppose the $n - r$ points $p_{r+1}, p_{r+2}, \ldots, p_n$ lies on the line s, and the r points $p_1, p_2, \ldots, p_r$ do not lie on s. Two lines p_ap_b and p_cp_d $(1 < a, c < r;\ r + 1 < b, d < n)$ are certainly distinct if $b \neq d$. Hence, among the $r(n - r)$ connecting lines p_ip_j $(i = 1, 2, \ldots, r;\ j = r + 1, r + 2, \ldots, n)$ at least

$$r(n - r) - \tfrac{1}{2}r(r - 1)$$

are distinct. Counting the line s we have

$$t > 1 + r(n - r) - \tfrac{1}{2}r(r - 1) = rn - \tfrac{1}{2}(3r + 2)(r - 1).$$

We now proceed to the proof of Theorem 4.1 and consider two cases.

Case 1.

$$\sum_{i=2}^{3k-1} v_i > 2.$$

In this case there are two points, say p_1 and p_2, each of which lies on at most $3k - 1$ connecting lines. Let s be the line through p_1 and p_2. The connecting lines through p_1 and p_2 other than s intersect in at most $(3k - 2)^2$ points. Hence, s contains at least $n - (3k - 2)^2$ points. Suppose it contains $n - x$ points, where

$$k < x < (3k - 2)^2.$$

Inequalities 4.70 and 4.71 insure that $n > \frac{3}{2}x$; hence, by the lemma, at least

$$xn - \tfrac{1}{2}(3x + 2)(x - 1)$$

connecting lines are determined. Using 4.70 and 4.71 we have

$$n \geq \tfrac{1}{2}\{3(x+k) - 1\}$$

or

$$n(x-k) \geq \tfrac{1}{2}(3x^2 - 3k^2 - x + k)$$

or

$$t \geq xn - \tfrac{1}{2}(3x+2)(x-1) \geq kn - \tfrac{1}{2}(3k+2)(k-1).$$

Case 2.

$$\sum_{i=2}^{3k-1} v_i < 1.$$

From 4.62 we have

$$3t \geq 3 + 2 + 3k(n-1)$$

or

$$t \geq kn - k + \frac{5}{3} > kn - \tfrac{1}{2}(3k+2)(k-1).$$

This completes the proof of Theorem 4.1.

Taking $k = 2$, we have

COROLLARY 4.1. *If at most $n - 2$ points are collinear and $n \geq 27$, then at least $2n - 4$ connecting lines are determined.*

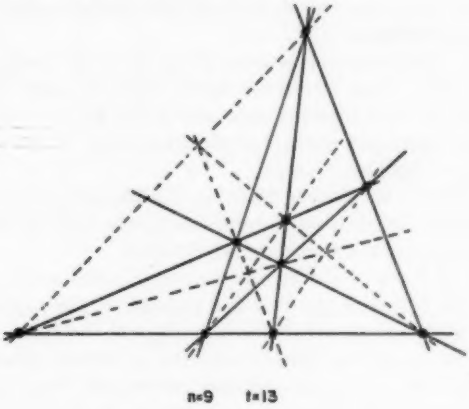

FIGURE 4.1

Figures 3.1, 3.2, and 4.1 show configurations with $n = 7, 8, 9$ and $t = 9, 11, 13$ respectively. On the other hand, using the above methods a somewhat detailed analysis leads to the conclusion that these are best possible, that is, if no $n - 1$ points are collinear and $n = 7, 8, 9$, then $t \geq 2n - 5$. Similarly, a detailed analysis shows that if $n = 10$ and no 9 points are collinear, then

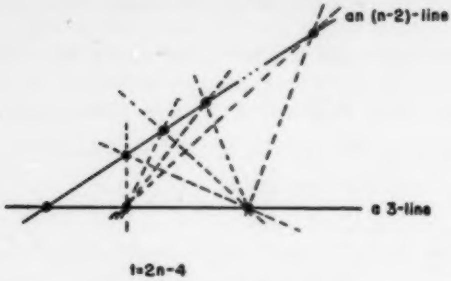

FIGURE 4.2

$t \geqslant 16$. Figure 4.2 shows a configuration with arbitrary n and $t = 2n - 4$. Thus, it seems very likely that, if at most $n - 2$ points are collinear, then $t \geqslant 2n - 5$ ($n = 7, 8, 9$) and $t \geqslant 2n - 4$ ($n > 9$) and that for each n there is a configuration for which equality holds.

The related problem of finding configurations for which t_3 is as large as possible is considered by Ball (**1**, pp. 105-6).

5. Zonohedra. In this section we point out a connection between the configurations we have been studying and the convex solids (in Euclidean space) known as zonohedra.

A *zonohedron* is a convex polyhedron whose faces all possess central symmetry (**2**, pp. 27-30). These properties insure that the solid has central symmetry. Each edge of a zonohedron determines a zone of faces in which each face has two sides equal and parallel to the given edge. If the edges occur in n different directions, there are n zones.

Let us call a set of n concurrent lines (in Euclidean space) a star. Then we may say that every zonohedron determines a star having one line parallel to each of the n directions in which the edges occur.

To every pair of faces of the zonohedra there corresponds the connecting plane (through the vertex of the star and parallel to the pair of faces) which contains the lines of the star parallel to the edges of the faces. The projection of the star and its connecting planes onto the projective plane at infinity is precisely a configuration of n non-collinear points and their connecting lines. Thus, a pair of parallel $2k$-gons on a zonohedron corresponds to a k-line of S. Theorem 3.6 now shows that:

Every zonohedron with n zones has at least $3n/7$ pairs of parallelogram faces.

References

1. W. W. R. Ball, *Mathematical Recreations and Essays* (11th ed., London, 1956).
2. H. S. M. Coxeter, *Regular Polytopes* (London, 1948).
3. ———, *A problem of collinear points*, Amer. Math. Monthly, *55* (1948), 26–28.
4. ———, *The Real Projective Plane* (2nd ed., Cambridge, 1955).
5. N. G. de Bruijn and P. Erdös, *On a combinatorial problem*, Hederl. Adad. Wetenach., *51* (1948), 1277–1279.
6. G. A. Dirac, *Collinearity properties of sets of points*, Quart. J. Math., *2* (1951), 221–227.
7. P. Erdös, *Problem No. 4065*, Amer. Math. Monthly, *51* (1944), 169.
8. ———, *On some geometrical problems*, Matematikai Lapok, *8* (1957), 86–92.
9. H. Hanani, *On the number of lines determined by n points*, Technion. Israel Inst. Tech. Sci. Publ., *6* (1951), 58–63 (Math. Rev., *17*, 294).
10. ———, *On the number of lines and planes determined by n points*, Technion. Israel Inst. Tech. Sci. Publ., *6* (1954/55), 58–63 (Math. Rev., *17*, 294).
11. D. W. Lang, *The dual of a well-known theorem*, Math. Gazette, *39* (1955), 314.
12. Th. Motzkin, *The lines and planes connecting the points of a finite set*, Trans. Amer. Math. Soc., *70* (1951), 451–464.
13. J. J. Sylvester, *Mathematical Question 11851*, Educational Times, *59* (1893), 98.
14. O. Veblen and J. W. Young, *Projective Geometry*, vol. 2 (Boston, 1918).

Michigan State University
and
University of Saskatchewan

ON THE NUMBER OF SIDES OF A PETRIE POLYGON

ROBERT STEINBERG

Let $\{p, q, r\}$ be the regular 4-dimensional polytope for which each face is a $\{p, q\}$ and each vertex figure is a $\{q, r\}$, where $\{p, q\}$, for example, is the regular polyhedron with p-gonal faces, q at each vertex. A Petrie polygon of $\{p, q\}$ is a skew polygon made up of edges of $\{p, q\}$ such that every two consecutive sides belong to the same face; but no three consecutive sides do. Then a Petrie polygon of $\{p, q, r\}$ is defined by the property that every three consecutive sides belong to a Petrie polygon of a bounding $\{p, q\}$, but no four do. Let $h_{p,q,r}$ be the number of sides of such a polygon, and $g_{p,q,r}$ the order of the group of symmetries of $\{p, q, r\}$. Our purpose here is to prove the following formula:

$$\frac{h_{p,q,r}}{g_{p,q,r}} = \frac{1}{64}\left(12 - p - 2q - r + \frac{4}{p} + \frac{4}{r}\right). \tag{1}$$

We use the following result of Coxeter (**1**, p. 232; **2**):

$$\frac{h_{p,q,r}}{g_{p,q,r}} = \frac{1}{16}\left(\frac{6}{h_{p,q} + 2} + \frac{6}{h_{q,r} + 2} + \frac{1}{p} + \frac{1}{r} - 2\right), \tag{2}$$

where $h_{p,q}$, for example, denotes the number of sides of a Petrie polygon of $\{p, q\}$. Both proofs referred to depend on the fact that the number of hyperplanes of symmetry of $\{p, q, r\}$ is $2h_{p,q,r}$. This is proved in a more general form in (**3**). Clearly (1) is a consequence of (2) and the following result:

If h is the number of sides of a Petrie polygon of the polyhedron $\{p, q\}$, then

$$h + 2 = \frac{24}{10 - p - q}. \tag{3}$$

Proof of (3). The planes of symmetry of $\{p, q\}$ divide a concentric sphere into congruent spherical triangles each of which is a fundamental region for the group $\mathfrak{G}$ of symmetries of $\{p, q\}$ (**1**, p. 81). The number of triangles is thus g, the order of $\mathfrak{G}$. The vertices of one of these triangles can be labelled P, Q, R so that the corresponding angles are π/p, π/q, $\pi/2$. There are $g/2p$ images of P under $\mathfrak{G}$, since the subgroup leaving P fixed has order $2p$. At each of these points there are $p(p - 1)/2$ intersections of pairs of circles of symmetry. Counting intersections at the images of Q and R in a similar fashion, one gets for the total number of intersections of pairs of circles of symmetry the number

Received October 21, 1957.

$g(p + q - 1)/4$. However, the number of such circles is $3h/2$ (**1**, p. 68), and every two intersect in two points. Hence

$$\frac{g(p + q - 1)}{4} = \frac{3h}{2}\left(\frac{3h}{2} - 1\right). \tag{4}$$

Dividing (4) by the relation $g = h(h + 2)$ of Coxeter (**1**, p. 91), and solving for h, one obtains (3).

References

1. H. S. M. Coxeter, *Regular polytopes* (London, 1948).
2. ———, *The product of the generators of a finite group generated by reflections*, Duke Math. J. *18* (1951), 765–782.
3. R. Steinberg, *Finite reflection groups*, submitted to Trans. Amer. Math. Soc.

University of California

CENTRAL LIMIT THEOREMS FOR INTERCHANGEABLE PROCESSES

J. R. BLUM, H. CHERNOFF, M. ROSENBLATT, AND H. TEICHER*

1. Introduction and summary. Let $\{X_n\}$ $(n = 1, 2, \ldots)$ be a stochastic process. The random variables comprising it or the process itself will be said to be interchangeable if, for any choice of distinct positive integers $i_1, i_2, i_3, \ldots, i_k$, the joint distribution of

$$X_{i_1}, X_{i_2}, \ldots, X_{i_k}$$

depends merely on k and is independent of the integers $i_1, i_2, \ldots, i_k$. It was shown by De Finetti **(3)** that the probability measure for any interchangeable process is a mixture of probability measures of processes each consisting of independent and identically distributed random variables. More precisely, let $\mathfrak{F}$ be the class of one-dimensional distribution functions and for each pair of real numbers x and y let

$$\mathfrak{F}(x, y) = \{F \in \mathfrak{F} | F(x) < y\}.$$

Let $\mathfrak{A}$ be the Borel field of subsets of $\mathfrak{F}$ generated by the class of sets $\mathfrak{F}(x, y)$. Then De Finetti's theorem asserts that for any interchangeable process $\{X_n\}$ there exists a probability measure μ defined on $\mathfrak{A}$ such that

$$P\{B\} = \int_{\mathfrak{F}} P_F\{B\} d\mu\,(F) \tag{1.1}$$

for any Borel measurable set B defined on the sample space of the sequence $\{X_n\}$. Here $P\{B\}$ is the probability of the event B and $P_F\{B\}$ is the probability of the event B computed under the assumption that the random variables X_n are independently distributed with common distribution function F.

Note that for Borel measurable point functions f for which the functional

$$\int_{-\infty}^{\infty} f(x) dF(x)$$

is well-defined,

$$\int_{-\infty}^{\infty} f(x) dF(x)$$

*Received June 26, 1957. Initially this problem was investigated by two pairs of authors: Blum and Rosenblatt on the one hand, and Chernoff and Teicher on the other. In view of the great overlap in results, a single combined paper was arranged.

Professor Blum's work was sponsored by the Office of Ordnance Research, U.S. Army under Contract DA-33-008-ORD-965; that of the other three authors was sponsored by the Office of Naval Research.

is measurable in F. This follows from the fact that it is true for f's that are indicators of half-lines. We then have for any integrable function g on the sample space of the sequence $\{X_n\}$

$$E\{g\} = \int E_F\{g\} d\mu(F)$$

where $E_F\{g\}$ is the expectation of g computed under the assumption that the random variables $\{X_n\}$ are independently distributed with common distribution function F.

In this paper we shall deal only with interchangeable processes having finite first and second moments and consequently shall assume without loss of generality that all such processes have mean zero and variance one. Let $\{X_n\}$ be such a process and for each positive integer n define

$$S_n = \sum_{i=1}^{n} X_i.$$

We shall say that the Central Limit Theorem holds for the process $\{X_n\}$ if for every real number α we have

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \phi(\alpha),$$

where

$$\phi(\alpha) = \frac{1}{\sqrt{[2\pi]}} \int_{-\infty}^{\alpha} e^{-\frac{1}{2}u^2}\, du.$$

In section two, necessary and sufficient conditions for the Central Limit Theorem to hold for an interchangeable process are derived. In section three, we discuss briefly the case of a doubly infinite sequence $\{X_{ni}\}$ where for each n the random variables X_{ni} are an interchangeable process. A number of conditions sufficient for the asymptotic normality of $S_n/\sqrt{n}$, where

$$S_n = \sum_{i=1}^{n} X_{ni}$$

are obtained.

2. The Single Process. Let X_n be an interchangeable process. According to (1.1) we have for every positive integer n

$$P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \int_{\mathfrak{F}} P_F\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} d\mu(F). \tag{2.1}$$

For each $F \in \mathfrak{F}$ define $m(F)$ and $\sigma(F)$ by

$$m(F) = \int_{-\infty}^{\infty} x dF(x) \text{ and } \sigma^2(F) = \int_{-\infty}^{\infty} [x - m(F)]^2 dF(x)$$

provided these integrals converge. For every real number m and non-negative number σ let $\mathfrak{F}_{m,\sigma}$ be the set of F for which $m(F) = m$ and $\sigma(F) = \sigma$. It can easily be shown that each such $\mathfrak{F}_{m,\sigma}$ is $\mathfrak{A}$-measurable. Now suppose $F \in \mathfrak{F}_{0,1}$.

Then it follows from the Central Limit Theorem for a sequence of independent and identically distributed random variable that

$$\lim_{n\to\infty} P_F\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \phi(\alpha).$$

Consequently, we see from (2.1) and the Lebesgue bounded convergence theorem that the process $\{X_n\}$ will satisfy the Central Limit Theorem if $\mu(\mathfrak{F}_{o,1}) = 1$. We shall show that this condition is also necessary. To do this, let $\mathfrak{F}'$ be the subset of $\mathfrak{F}$ for which $m(F)$ and $\sigma(F)$ exist and are finite. Again, it is easily seen that $\mathfrak{F}'$ is $\mathfrak{A}$-measurable and from the existence of the first and second moments of the process $\{X_n\}$ it follows that $\mu(\mathfrak{F}') = 1$. An easy computation shows that

$$\lim_{n\to\infty} P_F\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\}$$

exists for each $F \in \mathfrak{F}'$ and depends only on $m(F)$, $\sigma(F)$, and α. If we denote the limiting function by $f[m(F), \sigma(F), \alpha]$ we find

$$(2.2)\qquad f[m(F), \sigma(F), \alpha] = \begin{cases} 0 \text{ if } m(F) > 0, \\ 1 \text{ if } m(F) < 0, \\ 0 \text{ if } m(F) = 0, \sigma(F) = 0, \alpha < 0, \\ 1 \text{ if } m(F) = 0, \sigma(F) = 0, \alpha > 0, \\ \phi\left(\dfrac{\alpha}{\sigma(F)}\right) \text{ if } m(F) = 0, \sigma(F) > 0. \end{cases}$$

Also, let $\mathfrak{F}_o$ be the set of $F \in \mathfrak{F}$ for which $m(F) = 0$ and $\mathfrak{F}_{o,+}$ $(\mathfrak{F}_{o,o})$ be the set of $F \in \mathfrak{F}_o$ for which $\sigma(F) > 0$, $(\sigma(F) = 0)$. Now, if we again employ (2.1) and the Lebesgue bounded convergence theorem we find that if the Central Limit Theorem is to hold we must have

$$(2.3)\qquad \phi(\alpha) = \int_{\mathfrak{F}'} f[m(F), \sigma(F), \alpha]\, d\mu(F).$$

Let $\alpha > 0$. If we use (2.2), (2.3) and the fact that $\phi(-a) = 1 - \phi(\alpha)$ we find that

$$(2.4)\qquad 2\phi(\alpha) - 1 = \mu(\mathfrak{F}_{o,o}) + \int_{\mathfrak{F}_{o,+}} \left(2\phi\left[\frac{\alpha}{\sigma(F)}\right] - 1\right) d\mu(F).$$

On letting α approach infinity in (2.4) we have $\mu(\mathfrak{F}_o) = 1$.

Now let $G(\sigma)$ be the distribution function on the real line defined by

$$G(\sigma) = \begin{cases} 0 & \text{for } \sigma < 0 \\ \mu(F \mid m(F) = 0, \sigma(F) < \sigma) & \text{for } \sigma > 0. \end{cases}$$

Then we may write (2.3) in the form

$$(2.5)\qquad \phi(\alpha) = \int_o^\infty f[o, \sigma, \alpha] dG(\sigma).$$

If we put $\alpha = 0$ in (2.5) and use (2.2) we see that $G(0) = 0$. Thus we have

$$\phi(\alpha) = \int_{0+}^{\infty} \phi\left(\frac{\alpha}{\sigma}\right) dG(\sigma) = \frac{1}{\sqrt{[2\pi]}} \int_{0}^{\infty} \int_{-\infty}^{\alpha/\sigma} e^{-\frac{1}{2}u^2}\, du\, dG(\sigma). \tag{2.6}$$

Differentiating both sides of (2.6) with respect to α and setting $\alpha = 1$, we have

$$e^{-\frac{1}{2}} = \int_{0}^{\infty} \frac{1}{\sigma} e^{-\frac{1}{2}\sigma^2}\, dG(\sigma). \tag{2.7}$$

But the integrand of the right hand side of (2.7) achieves a unique maximum at $\sigma = 1$ where its value is $e^{-\frac{1}{2}}$. Thus, we see that (2.7) holds if and only if $G(\sigma)$ has all of its mass concentrated at the point $\sigma = 1$.

We summarize in

LEMMA 1. *If $\{X_n\}$ is an interchangeable process with mean zero and variance one the Central Limit Theorem holds if and only if $\mu(\mathfrak{F}_{0,1}) = 1$.*

The condition of the lemma is not very practical since in general it is rather difficult to compute the measure μ associated with a given interchangeable process $\{X_n\}$. However, we shall show that the condition of Lemma 1 is equivalent to a simple condition on the moments of the process. Suppose then that the condition of Lemma 1 holds. Then we have for $i \neq j$

$$\begin{aligned} E\{X_iX_j\} &= \int_{\mathfrak{F}_{0,1}} m^2(F)d\mu(F) = 0, \\ E\{X_i^2X_j^2\} &= \int_{\mathfrak{F}_{0,1}} [E_F\{X^2\}]^2 d\mu(F) = 1. \end{aligned} \tag{2.8}$$

Conversely, suppose (2.8) holds for $i \neq j$. Then $E\{[X_i^2 - 1][X_j^2 - 1]\} = 0$ and we obtain

$$\begin{aligned} &\int_{\mathfrak{F}} m^2(F)d\mu(F) = 0, \\ &\int_{\mathfrak{F}} [E_F\{X^2 - 1\}]^2 d\mu(F) = 0. \end{aligned} \tag{2.9}$$

But clearly (2.9) implies that $\mu(\mathfrak{F}_{0,1}) = 1$. Thus, we have

THEOREM 1. *Let $\{X_n\}$ be an interchangeable process with mean zero and variance one. Then the Central Limit Theorem holds for the process if and only if for $i \neq j$*

$$E\{X_iX_j\} = 0 \quad \text{and} \quad E\{[X_i^2 - 1][X_j^2 - 1]\} = 0.$$

We can rephrase the conditions of the theorem by saying that X_i and X_j as well as X_i^2 and X_j^2 must have covariance zero (be uncorrelated) for $i \neq j$.

Several remarks of interest can be made concerning such processes. In the first place, let $\{X_n\}$ be a sequence of independent and identically distributed random variables with mean m, variance σ^2, and finite third moment. Then, it was shown by Berry (1) and Esseen (4) that

$$\left| P\left\{ \frac{S_n - nm}{\sqrt{[n\sigma^2]}} < \alpha \right\} - \phi(\alpha) \right| < \frac{c}{\sqrt{n}} \frac{E\{|X - m|^3\}}{\sigma^3} \tag{2.10}$$

where c is a universal constant. It is simple to verify that if the Central Limit Theorem holds for an interchangeable process $\{X_n\}$ with finite third moment, then the Berry-Esseen bound still applies.

Secondly, consider an interchangeable process which is generated by a mixture over a family of one-dimensional distributions with the property that each distribution in the family is completely determined by specifying its mean and variance. The Normal distributions, the Poisson distributions and the Binomial distributions furnish examples of such families. But in such a case it follows easily from Lemma 1 that if the Central Limit Theorem holds, the mixture must be concentrated at a single distribution of the family. Consequently, we find that if the Central Limit Theorem holds for such a process, the random variables must be independent and identically distributed.

We observe that if $\{X_n\}$ $(n = 1, 2, \ldots)$ is an interchangeable process with $EX_1 = 0$, $EX_1^2 = 1$, $EX_1X_2 = \rho$ (necessarily non-negative) and such that every finite subset has a joint normal distribution,

$$\sum_{i=1}^{n} X_i$$

will be normal with mean 0 and variance $n + n(n - 1)\rho$. From Theorem 1 the normalization $1/\sqrt{n}$ will suffice only if $\rho = 0$ (the X_n are independent). However, if $\rho > 0$,

$$\frac{1}{n} \sum_{i=1}^{n} X_i$$

has a limiting normal distribution with mean 0 and variance ρ.

Finally, we consider again a sequence $\{X_n\}$ of independent and identically distributed random variables. Then, if $f(x)$ is any bounded measurable function and the sequence $\{Y_n\}$ is defined by $Y_n = f(X_n)$ it follows that the process $\{Y_n\}$ satisfies the Central Limit Theorem. However this is not, in general, true for interchangeable processes. For suppose $f(X)$ and $g(X)$ are bounded measurable functions and let $h(X) = f(X) + g(X)$. Then, if the Central Limit Theorem is to hold for the process $h(X_n)$ we find from Theorem 1 that we must have

$$E\{h(X_i)\, h(X_j)\} = [E\{h(X)\}]^2, \qquad i \neq j. \tag{2.11}$$

Now, from the interchangeability of the process $\{X_n\}$ it follows that

$$E\{f(X_i)\, g(X_j)\} = E\{f(X_j)\, g(X_i)\}$$

for all i and j. Using this we can expand the left side of (2.11) to obtain that

$$E\{f(X_i)\, g(X_j)\} = E\{f(X_i)\}\, E\{g(X_j)\}.$$

Since f and g are arbitrary, it follows that the random variables X_n are independently distributed.

3. Sequences of interchangeable processes. For each positive integer n, let $\{X_{ni}, i = 1, 2, \ldots, \}$ be an interchangeable process with mean zero, variance one, and finite absolute third moment. If we let μ_n denote the measure on $\mathfrak{A}$ occurring in the representation (1.1), it is clear that we must have $\mu_n\ (F|E_F\{|X|^3\} < \infty) = 1$ for every positive integer n.

By techniques paralleling those employed in the previous section we obtain

LEMMA 2. *Suppose that for every $\epsilon > 0$,*

$$\text{(i)} \qquad \lim_{n\to\infty} \mu_n\left(F|\ |m(F)| < \frac{\epsilon}{\sqrt{n}}\right) = 1,$$

$$\text{(ii)} \qquad \lim_{n\to\infty} \mu_n(F|\ |\sigma(F) - 1| < \epsilon) = 1,$$

$$\text{(iii)} \qquad \lim_{n\to\infty} \mu_n\left(F|\ \frac{E_F\{|X - m(F)|^3\}}{\sigma^3(F)} < \epsilon\sqrt{n}\right) = 1.$$

Then for every real number α

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \phi(\alpha).$$

For each integer n let $E_n\{\quad\}$ stand for the expectation of the quantity between the brackets computed with respect to the distribution of the nth process. With this notation we have

THEOREM 2. *For each positive integer n let $\{X_{ni}; i = 1, 2, \ldots\}$ be an interchangeable process with mean zero, variance one, and finite absolute third moment. If*

$$\text{(i)} \qquad E_n\{X_{n1}X_{n2}\} = o\left(\frac{1}{n}\right),$$

$$\text{(ii)} \qquad \lim_{n\to\infty} E_n\{X_{n1}^2X_{n2}^2\} = 1,$$

$$\text{(iii)} \qquad E_n\{|X_{n1}|^3\} = o(\sqrt{n}),$$

then for every real number α

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \phi(\alpha).$$

The theorem is obtained by showing that conditions (i), (ii), and (iii) of Lemma 2 are satisfied.

The conditions of Theorem 2 are not necessary. However, it is of some interest to remark that in a certain sense these conditions are the best of their kind. Given any one of the three conditions, one can find an interchangeable process with mean zero, unit variance, and finite third moment which narrowly violates the condition, satisfies the remaining two conditions and for which the Central Limit Theorem is not valid.

A somewhat different limit theorem can be obtained in the following way. For each positive integer n let $\{X_{ni}\}$ be an interchangeable process with mean zero, unit variance, and mixing measure μ_n. Define the distribution function $F_n(m)$ by

$$F_n(m) = \mu_n\{F \mid \sqrt{n}\, m(F) < m\} \text{ for } n = 1, 2, \ldots .$$

Let $F(m)$ be an arbitrary distribution. Then we have

THEOREM 3. *If*

$$\text{(i)} \qquad \lim_{n\to\infty} F_n(m) = F(m)$$

at every continuity point m *of* F,

$$\text{(ii)} \qquad \lim_{n\to\infty} E_n\{X_{n,1}X_{n,2}\} = 1,$$

and

$$\text{(iii)} \qquad E_n\{|X_{n,1}|^3\} = o(\sqrt{n}),$$

then

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = \int_{-\infty}^{\infty} \phi(\alpha - m)\,dF(m), \qquad -\infty < \alpha < \infty.$$

The result is obtained by making use of the Helly-Bray theorem.

We note that the limiting distribution obtained in Theorem 3 is the convolution of ϕ and F. Consequently, it may be regarded as the distribution of the sum of two independent random variables, one of which is normal with mean zero and variance one and the other with distribution F. It follows from a theorem of Cramer **(2)** that the limit distribution is normal if and only if $F(m)$ is a normal distribution. Now, let $N_{a,b}(\alpha)$ be the normal distribution with mean a and variance b, that is,

$$N_{a,b}(\alpha) = \frac{1}{\sqrt{[2\pi b]}} \int_{-\infty}^{\alpha} e^{-(x-a)^2/2b}\, dx.$$

Further, for $k = 1, 2, \ldots$, let a_k denote the kth moment of

$$N_{a_1,b}(\alpha).$$

Then we can give a criterion for normality of the limiting distribution in terms of moment conditions on the process as follows:

COROLLARY. *If condition* (i) *in Theorem* 3 *is replaced by the condition*

$$\lim_{n\to\infty} E_n\{n^{k/2}X_{n1}X_{n2}\ldots X_{nk}\} = a_k \qquad k = 1, 2, \ldots$$

then

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} < \alpha\right\} = N_{a_1,1+b}(\alpha).$$

References

1. A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. Amer. Math. Soc., *49* (1941), 122–136.
2. Harold Cramer, *Random variables and probability distributions*, Cambridge Tracts in Mathematics, *136* (Cambridge, 1937).
3. Bruno De Finetti, *La prévision, ses lois logiques, ses sources subjectives*, Annales de l'Institut Henri Poincaré, 7 (1937), 1–68.
4. Carl-Gustav Esseen, *Fourier analysis of distribution functions*, Acta Math., 77 (1944), 1–125.

Indiana University
Stanford University
Purdue University

SOME GENERALIZATIONS OF THE PROBLEM OF DISTINCT REPRESENTATIVES

N. S. MENDELSOHN AND A. L. DULMAGE

1. Introduction. If $S_1, S_2, S_3, \ldots, S_n$ are subsets of a set M then it is known that a necessary and sufficient condition that it is possible to choose representatives a_i such that a_i is in S_i for $(i = 1, 2, 3, \ldots, n)$ and such that $a_i \neq a_j$ for $i \neq j$, is that for $k = 1, 2, 3, \ldots, n$, the union of any k of the sets $S_1, S_2, \ldots, S_n$, contains at least k elements. The theorem has a number of consequences amongst which we list the following.

(1) If a set M containing rs elements be broken up in two different ways into r disjoint subsets each containing s elements then it is possible to find elements $a_1, a_2, \ldots, a_r$ which will serve as representatives of the sets of both decompositions **(7)**, **(15)**.

(2) If A is an r by r matrix whose entries are all either one or zero and if each row and each column contains exactly s ones then A can be expressed as a sum of s permutation matrices **(15)**.

(3) An n by n Latin square can always be completed when m of its rows $(m < n)$ are specified **(5)**.

(4) Any doubly stochastic matrix A lies in the convex closure of the permutation matrices. More particularly, any doubly stochastic matrix is a weighted average of at most $n^2 - n + 1$ permutation matrices **(2)**.

Ryser and Mann **(15)** generalized the representative problem to obtain sufficient conditions that specified elements $a_1, a_2, a_3, \ldots, a_r$ may appear in a system of distinct representatives and Hoffman and Kuhn **(9)** replaced these by conditions which are both necessary and sufficient. These generalized results can be used to prove a theorem due to Ryser **(18)** which states necessary and sufficient conditions in order that a specified r by s subrectangle be completable to an n by n latin square.

In this paper we obtain a simple combinatorial proof of a generalization of the theorem of Hoffman and Kuhn. We also obtain generalizations in other directions which enable us to extend some of the results enumerated above.

For our purpose it is more convenient to use an alternative equivalent formulation of the distinct representative theorem. We first define a few terms.

Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_s$ $(n < s)$ be two sets of elements. Let R be a dyadic relation connecting an a with a b. A pair (a, b) will be called an incidence if the relation R holds for a and b. A set S of incidences

$$(a_{i_1}, b_{r_1}),\ (a_{i_2}, b_{r_2}),\ \ldots\ (a_{i_k}, b_{r_k})$$

Received June 21, 1957.

will be called regular if no a or b appears more than once amongst the components of the pairs of S. The distinct representative theorem can now be reworded as follows:

A necessary and sufficient condition that a regular set of incidences which includes all of the $a_1, a_2, \ldots, a_n$ exist, is that for each k, $(k = 1, 2, 3, \ldots, n)$ every subset of k of the a_i are incident with at least k distinct b_j.

In connection with the sets $\{a_i\}$, $\{b_j\}$ and the incidence relation R, we define an incidence matrix A to be an n by s matrix whose entry in the ith row and jth column is 1 if (a_i, b_j) is an incidence and 0 otherwise.

An m by p matrix will be called a sub permutation matrix of rank r if it satisfies the following conditions:

(1) Its entries are all 0 or 1.

(2) Each row and each column contains at most one 1.

(3) The matrix contains exactly r 1's.

The set of places at which the 1's appear in a sub permutation matrix of rank r will be called a sub permutation set of places of rank r.

With this new notation the distinct representative theorem simply states that if the stated conditions on the a's and b's are satisfied the incidence matrix will contain 1's at a sub permutation set of places of rank n.

2. A generalization of the Hoffman-Kuhn Theorem. The following theorem generalizes the Hoffman-Kuhn theorem which was given in (9). It has the advantage that the result is completely symmetric in the a's and b's.

THEOREM 1. *Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_m$ be two sets of elements connected by an incidence relation R. A necessary and sufficient condition that a regular set S of incidences exist in which $a_1, a_2, \ldots, a_r$ and $b_1, b_2, \ldots, b_s$ appear is:*

(1) *For $k = 1, 2, \ldots, r$, any subset of k of the elements of $a_1, a_2, \ldots, a_r$ are incident with at least k distinct elements of $b_1, b_2, \ldots, b_m$.*

(2) *For $p = 1, 2, \ldots, s$, any subset of p of the elements $b_1, b_2, \ldots, b_s$ are incident with at least p distinct elements of $a_1, a_2, \ldots, a_n$.*

Proof. Form the incidence matrix A. By condition (1), using the distinct representative theorem, the first r rows of A contain at least one sub permutation matrix of rank r. Put the letter R in each of the places occupied by 1 in any one such sub permutation matrix. Similarly the first s columns of A contain at least one sub permutation matrix of rank s. Put the letter C in the places occupied by 1 in any such sub permutation matrix. (It is possible for the same place to be marked with both R and C.) The matrix A is now said to be marked by a set of R-places and a set of C-places. It will now be shown how to choose a subset of the R and C places which will produce the set of incidences S required by the theorem.

(1) If a place is marked by both R and C it is to be included in the subset.

(2) If a marked element is alone in its row or column it is the beginning of

a connected chain which ends with an element which is alone in its row or column. A connected chain is defined as a sequence of marked places formed as follows. Start with any marked place which is alone in its row or column and proceed to a marked place in the same row (column), continue to a further marked place in the same column (row) and continue as far as possible. The set of marked places visited by this procedure is called a connected chain. By our definition a single marked element alone in its row and column is a chain. There are four types of connected chain.

Type 1. The chain begins and ends with an R. (See Figure 1.)

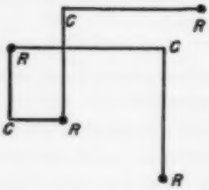

FIGURE 1

In this case we must include the R's and omit the C's in the required subset.

Type 2. The chain begins with an R and ends with a C and the first step is horizontal. (See Figure 2.)

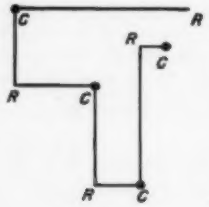

FIGURE 2

In this case we must include the C's and omit the R's from the required subset.

Type 3. The chain begins with an R and ends with a C and the first step is vertical. (See Figure 3.)

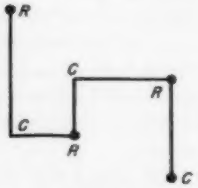

FIGURE 3

In this case the required subset must include the R's and omit the C's.

Type 4. The chain begins and ends with a C.

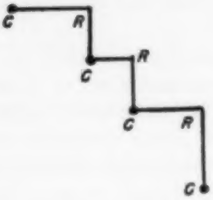

FIGURE 4

In this case the required subset must include the C's and omit the R's.

In view of the fact that the beginning and end of a chain are interchangeable, no other types of chain are possible.

(3) After removing from the set of marked places all doubly marked places and all chains, either there are no further marked places or the remaining marked places (which we will call the core) have the following property. Each marked place in the core has another marked place in its row and another marked place in its column. Hence, the marked places in the core all lie in a square sub-matrix of the incidence matrix which is included in the r by s sub-matrix of A which lies in the upper left corner. The required subset can now be completed by choosing from the core either all the places marked R or all the places marked C. (There are other ways of making the choice as our next theorem will show.)

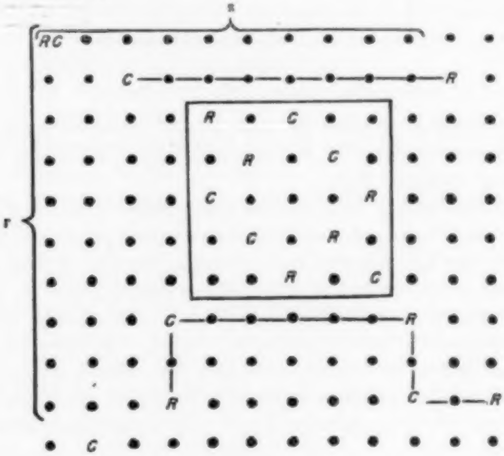

FIGURE 5

It is clear that the chosen subset of marked places yields a set of incidences S which satisfies the conditions of the theorem. In Figure 5 is shown an R and C marking of an incidence matrix. Here $m = 12$, $s = 10$, $n = 11$, $r = 10$. The chains are marked by lines and the core is surrounded by a singly lined square.

A question of some interest is the following. From an incidence matrix A marked by a set of R-places and a set of C-places in how many ways is it possible to choose a subset of places which yield a regular set S of incidences? From our proof of Theorem 1 it is immediate that the inclusion or exclusion of a place which is doubly marked or which belongs to a connected chain is uniquely determined. With regard to the core, its places lie in a square. By permuting the rows and the columns of this square it is possible to arrange that the marked squares can be confined to a set of square blocks along the main diagonal of the square as in Figure 6.

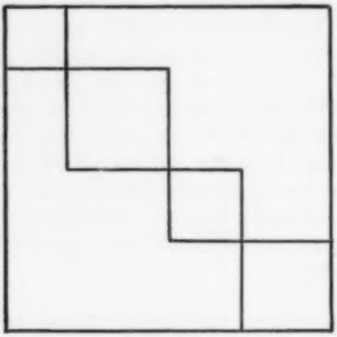

FIGURE 6

Let k be the maximal number of such blocks. In each such block we can choose for our subset either all the R's or all the C's. Hence, the number of ways of making a choice is 2^k.

There is another way of determining the number k which does not require the rearrangement of the rows and columns of the square containing the core. Start with any marked place of the core and proceed to the other marked place in its row, continue to the remaining marked place in the column of the place last visited and proceed in this way alternately along rows and columns. Ultimately the starting point is reached. The marked places visited in this manner form a re-entrant cycle and the number k described above is precisely the number of re-entrant cycles in the core. In Figure 7 there is a diagram of a core which is included in a 7 by 7 square and which decomposes into two cycles.

The theorem just proved can be described as follows.

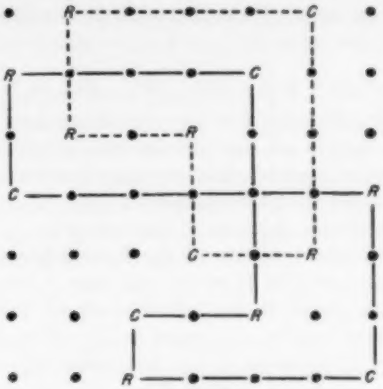

FIGURE 7

THEOREM 2. *The number of ways of choosing a set of places which yield a regular set of incidences from an R and C marking of an incidence matrix is 2^k, where k is the number of re-entrant cycles in the core of the marking.*

3. Further generalizations. We now consider generalizations of a different character. Roughly stated our problem is this. Let A be a matrix with entries which are either positive numbers or zero. What conditions can be placed on the entries of A in order to assure that A has non-zero entries in a sub permutation set of places or rank r? We confine ourselves to the case where A is a square matrix, since rectangular matrices may be completed into squares by adding rows (or columns) of zero entries. All resulting theorems for matrices so augmented will hold for the original matrix.

Throughout the remainder of this section the following notation will be used. $A = (a_{ij})$ will represent an n by n matrix with entries $a_{ij} \geqslant 0$. R_i will denote the sum of the entries in the ith row, $R_i = \sum_j a_{ij}$; C_j will denote the sum of the entries in the jth column, $C_j = \sum_i a_{ij}$; M will denote the maximum row or column sum, $M = \max(R_i, C_j)$; and S will denote the sum of all the entries,

$$S = \sum_{i,j} a_{ij} = \sum_i R_i = \sum_j C_j.$$

We will also use the following consequence of the distinct representative theorem: if A is a square matrix with positive or zero entries such that each row sum and each column sum has the same non-zero value, then any non-zero entry of A lies in a permutation set of places all of which have non-zero entries in A.

THEOREM 3. *If $(n - 1)\, M < S$ then A has non-zero entries in at least one permutation set of places.*

Proof. Augment the matrix A by adding an $(n + 1)$th row and an $(n + 1)$th column where $a_{i,\,n+1} = M - R_i$ $(i = 1, 2, \dots n)$

$$a_{n+1,\,j} = M - C_j (j = 1, 2, \dots n) \quad \text{and} \quad a_{n+1,n+1} = S - (n - 1)M.$$

Since all of $M - R_i$, $M - C_j$, $S - (n - 1)$ M are positive or zero and since in the augmented matrix all row and column sums are the same positive number, the augmented matrix has non-zero entries in a permutation set of places which includes the place occupied by $a_{n+1,\,n+1}$. The remaining places are thus a permutation set of places of the matrix A.

Theorem 3 is the best possible in the following sense: if the condition $(n - 1)M < S$ is not satisfied there are matrices A which do not have non-zero entries in all the places of any permutation set. Indeed, if A be any sub permutation matrix of rank $n - 1$, then $(n - 1)M = S$ and the theorem is obviously false for A. Theorem 3 has the following corollary which is an improvement of the result due to Hall **(7)**.

COROLLARY. *Let T be a set containing S elements and suppose T is broken up into n disjoint subsets in two different ways;*

$$T = A_1 + A_2 + \dots + A_n,$$
$$T = B_1 + B_2 + \dots + B_n.$$

Let M be the maximum number of elements in any of the sets $A_1, A_2, \dots, A_n$; $B_1, B_2, B_3 \dots, B_n$. If $(n - 1)M < S$ then it is possible to choose n elements $a_1, a_2, \dots, a_n$ which will represent both the sets $A_1, A_2, \dots, A_n$ and $B_1, B_2, B_3, \dots, B_n$ (each a_i being a member of the sets which it represents).

Proof. Form the matrix A whose entry a_{ij} is the number of elements in the intersection of A_i and B_j. $(i, j = 1, 2, \dots, n)$ A satisfies the condition of Theorem 3 and the corollary follows immediately.

THEOREM 4. *If*

$$\frac{1}{n-1} < \frac{M}{S} < \frac{n-1}{n^2 - 2n}, \tag{3.1}$$

then the matrix A has non-zero entries in the places of at least one sub permutation set of rank $(n - 1)$.

Proof. Let $B_2 = n\,M - S - M$ and $T_2 = M + B_2 - n\,B_2$. The inequality $S < (n - 1)M$ implies $B_2 > 0$ while the inequality

$$(n^2 - 2n)M < (n - 1)S$$

implies $T_2 > 0$. Augment the matrix A by the addition of two rows and two columns as follows. Put $a_{i,n+1} = M - R_i$, $a_{i,n+2} = B_2$ for $i = 1, 2, \dots, n$; put $a_{n+1,n+1} = a_{n+1,n+2} = a_{n+2,n+1} = 0$; $a_{n+2,n+2} = T_2$; $a_{n+1,j} = M - C_j$, $a_{n+1,j+1} = B_2$ for $j = 1, 2, \dots, n$. The augmented matrix now consists of non-negative entries whose row and column sums are all equal. Hence, since $a_{n+2,n+2} \neq 0$, there is a permutation set of places which includes the place $n + 2$, $n + 2$

containing non-zero elements of the augmented matrix. The places of this permutation set which are in the matrix A form a sub permutation set of rank at least $n - 1$.

The conditions of Theorem 4 are sufficient but may not be the best possible. Since sub permutation matrices of rank $n - 2$ satisfy the condition $(n - 2)M = S$, it is natural to conjecture that the term $(n - 1)/(n^2 - 2n)$ could be replaced by $1/(n - 2)$. For large n the improvement is small.

A corollary analogous to that of Theorem 3 reads as follows:

Let T be a set containing S elements and suppose T is broken up into n disjoint subsets in two different ways;

$$T = A_1 + A_2 + \dots + A_n = B_1 + B_2 + \dots + B_n.$$

Let M be the maximum number of elements in any of the sets $A_1, A_2, \dots, A_n$; $B_1, B_2, \dots, B_n$. If (3.1) holds then $(n - 1)$ elements may be chosen, each of which represents one set of each decomposition, and two elements may be chosen to represent the remaining two sets provided they are non-null.

We now proceed to the general case by induction. We now define numbers $B_2, B_3, B_4, \dots$; $T_1, T_2, T_3, T_4, \dots$ as follows:

$$\begin{aligned}
B_2 &= (n - 1)M - S,\\
B_3 &= n\,B_2 - B_2 - M,\\
B_4 &= n\,B_3 - B_2 - B_3 - M,\\
B_5 &= n\,B_4 - B_2 - B_3 - B_4 - M/\\
B_r &= n\,B_{r-1} - (B_2 + B_3 + \dots + B_{r-1}) - M,
\end{aligned}$$

and for $i = 1, 2, 3, \dots, T_i = -\mathrm{B}_{i+1}$. It is clear that if T_r is the first of the numbers $T_1, T_2, T_3, \dots$ which is positive, then all the numbers $B_2, B_3, \dots, B_r$ are positive.

The B_i may be expressed in terms of n, S, M as follows:

$$\begin{aligned}
B_2 &= (n - 1)M - S,\\
B_3 &= (n^2 - 2n)M - (n - 1)S,\\
B_i &= P_{i-1}(n)M - Q_{i-2}(n)S,
\end{aligned}$$

where $P_i(n)$ and $Q_i(n)$ of polynomials in n with integral coefficients of degree i.

By subtracting two successive B's one obtains $B_i' = n(B_{i-1} - B_{i-2})$. This in turn yields the following recurrence relationships amongst the P_i and the Q_i:

$$P_i = n(P_{i-1} - P_{i-2});\ Q_i = n(Q_{i-1} - Q_{i-2}).$$

Also, since $Q_0 = 1$, $Q_1 = (n - 1)$, $Q_2 = n^2 - 2n$; $P_1 = n - 1$, $P_2 = n^2 - 2n$, it follows that $P_i(n) = Q_i(n)$. Furthermore, the difference equation for P_i together with P_1 and P_2 uniquely determines P_i. It is directly verifiable by induction that

$$P_i(n) = n^i - \binom{i}{1} n^{i-1} + \binom{i-1}{2} n^{i-2} - \dots$$
$$= \sum_{k=0}^{[\frac{1}{2}i+\frac{1}{2}]} (-1)^k \binom{i-k+1}{k} n^{i-k}.$$

Before we can proceed to the main theorem we need some further properties of the $P_i(n)$.

THEOREM 5. *For $r > 2$ and $n \geqslant r$, $P_r(n) > 0$.*

Proof. We can bracket the terms of $P_r(n)$ as follows:

$$P_r(n) = \left\{n^r - \binom{r}{1} n^{r-1}\right\} + \left\{\binom{r-1}{2} n^{r-2} - \binom{r-2}{3} n^{r-3}\right\} + \dots$$
$$+ \left\{\binom{r-2k+1}{2k} n^{r-2k} - \binom{r-2k}{2k+1} n^{r-2k-1}\right\} + \dots$$

If each bracketed pair is non-negative for any positive value of n, it remains non-negative when n is replaced by a larger value. Hence, it is sufficient to prove that each bracketed pair of $P_r(r) \geqslant 0$ (and at least one pair has a value > 0). The theorem is trivially verified for $r = 3$ and $r = 4$ so we will assume $r \geqslant 5$.

Now

$$\binom{r-2k+1}{2k} r^{r-2k} - \binom{r-2k}{2k+1} r^{r-2k-1}$$
$$= r^{r-2k-1}\left\{\binom{r-2k+1}{2k} r - \binom{r-2k}{2k+1}\right\}.$$

The expression

$$\binom{r-2k+1}{2k} r - \binom{r-2k}{2k+1}$$

will be non-negative provided

$$(2k-1)r^2 - (4k^2 - 10k + 1)r > 4k(4k-1).$$

Here r is an integer $\geqslant 5$ and k is any integer such that $4k + 1 \leqslant r$. Again:

$$\begin{aligned}&(2k-1)r^2 - (4k^2 - 10k + 1)r\\ &\quad \geqslant r\{(2k-1)(4k+1) - (4k^2 - 10k + 1)\}\\ &\quad = r\{4k^2 + 8k - 2\} \geqslant 5\{4k^2 + 8k - 2\} > 4k(4k-1)\end{aligned}$$

for every positive integral k.

THEOREM 6. *For every integral $n > 2$,*

$$\frac{P_0(n)}{P_1(n)} < \frac{P_1(n)}{P_2(n)} < \frac{P_2(n)}{P_3(n)} \dots < \frac{P_{n-2}(n)}{P_{n-1}(n)} < \frac{P_{n-1}(n)}{P_n(n)}.$$

Proof. By Theorem 5 all the denominators are positive. Also, since $P_i = n(P_{i-1} - P_{i-2})$ for every i, it follows that

$$\begin{aligned}
P_r^2 - P_{r-1}P_{r+1} &= P_r^2 - P_{r-1}(n\,P_r - n\,P_{r-1}) \\
&= P_r^2 - n\,P_r\,P_{r-1} + n\,P_{r-1}^2 \\
&= (n\,P_{r-1} - n\,P_{r-2})\,P_r - n\,P_r\,P_{r-1} + n\,P_{r-1}^2 \\
&= n(P_{r-1}^2 - P_{r-2}\,P_r).
\end{aligned}$$

Also $P_1^2 - P_0\,P_2 = 1$, so that $P_r^2 - P_{r-1}\,P_{r+1} = n^{r-1} > 0$, for $r > 2$, $n \geqslant r$.

THEOREM 7. *Let A bt an n by n matrix with non-negative entries. Let S be the sum of all entries in A and let M be the maximum sum of any row or column of A. For $r > 2$ and $n \geqslant r$, if*

$$\frac{P_{r-2}(n)}{P_{r-1}(n)} \leqslant \frac{M}{S} < \frac{P_{r-1}(n)}{P_r(n)}, \tag{3.2}$$

then A has non-negative entries in a subpermutation set of places of rank $n - r + 1$.

Proof. By (3.2), $M\,P_r(n) - S\,P_{r-1}(n)$ is negative. That is, B_{r+1} is negative or T_r is positive. Also, since

$$\frac{P_0}{P_1} < \frac{P_1}{P_2} \cdots \frac{P_{r-2}}{P_{r-1}} \leqslant \frac{M}{S},$$

each of the numbers $B_2, B_3, B_4 \ldots B_r$ is non-negative. Augment the matrix A by the addition of r rows and r columns as follows: put

$$\begin{gathered}
a_{i,\,n+1} = M - R_i \text{ for } i = 1, 2, 3, \ldots, n; \\
a_{i,\,n+t} = B_t \text{ for } i = 1, 2, \ldots n \text{ and } t = 2, 3, \ldots, r; \\
a_{n+1,\,j} = M - C_j \text{ for } j = 1, 2, 3, \ldots, n; \\
a_{n+u,\,j} = B_u \text{ for } u = 2, 3, \ldots r \text{ and } j = 1, 2, \ldots, n; \\
a_{vs} = 0 \text{ for } v > n,\ s > n \text{ and } v \neq s \\
a_{n+k,\,n+k} = B_{k+2} + B_{k+3} + \ldots + B_r \text{ for } k = 1, 2, 3, \ldots, r - 2; \\
a_{n+r-1,\,n+r-1} = 0,\ a_{n+r,\,n+r} = T_r.
\end{gathered}$$

Figure 8 illustrates the case $r = 5$.

$$\begin{array}{ccccccccc}
\multicolumn{4}{c}{\multirow{3}{*}{\boxed{\qquad\qquad\qquad A \qquad\qquad\qquad}}} & M - R_1, & B_2, & B_3, & B_4, & B_5 \\
 & & & & M - R_2, & B_2, & B_3, & B_4, & B_5 \\
 & & & & M - R_n, & B_2, & B_3, & B_4, & B_5 \\
M - C_1, & M - C_2, & \ldots, & M - C_n, & B_3 + B_4 + B_5, & 0 & ,0 & ,0 & ,0 \\
B_2 & ,\ B_2 & ,\ldots, & B_2 & ,\ 0 & ,\ B_4 + B_5, & 0 & ,0 & ,0 \\
B_3 & ,\ B_2 & ,\ldots, & B_2 & ,\ 0 & ,\ 0 & ,\ B_5, & 0 & ,0 \\
B_4 & ,\ B_4 & ,\ldots, & B_4 & ,\ 0 & ,\ 0 & ,0 & ,0 & ,0 \\
B_5 & ,\ B_5 & ,\ldots, & B_5 & ,\ 0 & ,\ 0 & ,0 & ,0 & ,\ T_5
\end{array}$$

FIGURE 8

The augmented matrix has all its entries non-negative with the same non-zero sum for each row and column. Hence, it contains a permutation set of places which includes the place occupied by T_r. Those places which lie in A form a sub permutation set of places of rank at least $n - r + 1$.

4. Concluding Remarks.

The condition (3.2) for our main theorem, 7, while sufficient to assure the existence of a sub permutation set of places of non-zero entries of rank $n - r + 1$, may not be the best possible. Since a sub permutation matrix of rank k satisfies the condition $kM = S$, it seems reasonable to conjecture that the above condition may be replaced by the condition

$$\frac{1}{n - r + 1} < \frac{M}{S} < \frac{1}{n - r}.$$

If this latter condition is correct the result must be the best possible. For large n (that is, large in comparison with r) the difference between the proved result and the conjectured one is very small.

Note added in proof. (Feb. 20, 1958). The expression *term rank* has been used recently to describe the order of the largest minor of A with a non-zero term in the expansion of its determinant. In section 3 information concerning the term rank of a matrix A was obtained by embedding A in a doubly stochastic matrix. The nature and structure of such embedding has since been studied by the authors and the concept of *stochastic rank* of a matrix has been introduced as follows. An n by n matrix A with non-negative entries has stochastic rank σ if A can be embedded in a doubly stochastic matrix by the addition of $n - \sigma$ rows and columns but A cannot be embedded in a doubly stochastic matrix of smaller size. The following results concerning the stochastic rank σ and term rank ρ have been obtained.

(a) For any matrix A, $\rho \geqslant \sigma$.

(b) For a doubly stochastic matrix, or for a sub permutation matrix $\rho = \sigma$.

(c) For any matrix A, $\sigma = [S/M]$.

(d) There are n by n matrices A for which $\rho - \sigma = n - 1$.

(e) If S/M is not an integer $\rho \geqslant \sigma + 1$.

Furthermore, the conjecture stated in the concluding remarks has now been proved and extended to non-square and to infinite matrices.

These and other results will be proved in a sequel to the present paper.

References

1. G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Rev. Univ. Nacional de Tucuman, Ser. A, *5* (1946), 147–150.

2. L. Dulmage and I. Halperin, *On a theorem of Frobenius-König and J. von Neumann's game of hide and seek*, Trans. Roy. Soc. Can., Ser. III, *49* (1955), 23–29.

3. C. J. Everett and G. Whaples, *Representations of sequences of sets*, Amer. J. Math., *71* (1949), 287–293.

4. G. Frobenius, *Ueber matrizen mit nicht negativen elementen*, S. B. Berliner Akad., *23* (1912), 456–477.

5. M. Hall, *An existence theorem for latin squares*, Bull. Amer. Math. Soc., *51* (1945), 387–388.

6. ———, *Distinct representatives of subsets*, Bull. Amer. Math. Soc., *54* (1948), 922–926.

7. P. Hall, *On representatives of subsets*, J. Lond. Math. Soc., *10* (1935), 26–30.

8. P. R. Halmos and H. E. Vaughan, *The marriage problem*, Amer. J. Math., *72* (1950), 214–215.

9. A. J. Hoffman and H. W. Kuhn, *Systems of Distinct representatives and linear programming*, Amer. Math. Monthly, *63* (1956), 455–460.

10. ——— and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J., *20* (1953), 37–39.

11. D. König, *Ueber Graphen und ihre anwendung*, Math. Ann., 77 (1916), 453–465.

12. ———, *Theorie der endlichen und unendlichen graphen* (Chelsea, New York, 1950), 170–178.

13. W. Maak, *Eine neue definition der fastperiodicshen functionen*, Abh. Math. Sem. Hamb., *11* (1936), 240–244.

14. H. B. Mann, *Analysis and design of experiments* (Dover, New York, 1949).

15. ———, and H. J. Ryser, *Systems of distinct representatives*, Amer. Math. Monthly, *60* (1953), 397–401.

16. J. von Neumann, *A certain zero-sum two person game equivalent to the optimal assignment problem*, Contribution to the theory of games II, Ann. Math. Studies, no. 28 (1950), 5–12.

17. R. Rado, *Bemerkungen zur kombinatorik im anschluss an untersuchen von herrn D. König*, Setz. Ber. Math. Geselt, *32* (1933), 60–75.

18. H. J. Ryser, *A combinatorial theorem with an application to latin rectangles*, Proc. Amer. Math. Soc., *2* (1951), 550–552.

19. C. A. B. Smith and H. O. Hartley, *The construction of Youden squares*, J. Roy. Statist. Soc., Ser. B, *10* (1948), 262–263.

20. E. Sperner, *Note zu der arbeit von herrn B. L. von der Waerden: ein satz ueber klasseneinteilungen von endlichen mengen*, Abh. Math. Sem. Hamb., *5* (1926), 232.

21. B. L. von der Waerden, *Ein satz ueber klasseneinteilungen von endlichen mengen*, Abh. Math. Sem. Hamb., *5* (1926), 185–188.

22. H. Weyl, *Almost periodic invariant vector sets in a metric vector space*, Amer. J. Math., *71* (1949), 178–205.

ALGEBRAIC APPROXIMATION OF CURVES

A. H. WALLACE

Introduction. In his paper on the algebraic approximation of differentiable manifolds Nash **(1)** introduced the concept of a sheet of a real algebraic variety (see the definition in §16 below) and raised certain questions of a general nature. In attempting to answer these questions it has been necessary to evolve some sort of technique for manipulating curves on algebraic varieties, and, in particular, to set up a criterion for the possibility of approximating a sequence of analytic arcs (definition in §1) joined end to end by a single analytic arc. The greater part of this paper is devoted to this latter topic, the results being applied in the last section to the problems suggested in Nash's paper.

The work falls naturally into three parts. The first deals with the approximation of plane curves by algebraic curves, the second with the corresponding problem in higher dimensional spaces and on varieties in general, while the third is concerned with the sheets of real algebraic varieties. The separate preliminary treatment of the case of plane curves is natural in the sense that plane algebraic curves present a specially simple situation, being represented each by a single equation.

The following is a brief sketch of the paper. Part I deals with plane curves consisting of analytic arcs placed end to end, the object being to approximate these by parts of algebraic curves, smoothing off the joins of the arcs in some way but preserving in some way the other singularities of the original curve. The corresponding question in Euclidean n-space is then taken up in Part II. Finally, for a curve on a real algebraic variety the technique is to project into a suitable linear space, approximate the projected curve and then lift back on to the variety again. It is here that the preservation of the singularities of the given curve is important. For a bit of experimentation soon shows that the approximation of the projected curve may not lift into an approximation of the original curve unless attention is given to this point. A disturbance of the structure of a singularity may result in the lifted curve going off, so to speak, in the wrong direction.

In Part III the sheets of a real algebraic variety are defined, namely, as sets maximal with respect to the property that any two points can be joined by an analytic arc in the set. This property of analytic connectivity is not transitive, and so the concept is a bit tricky to handle, but the use of the approximation theorems of Part II gives a partial transitivity property. Next, a local study of a real algebraic variety shows that it has locally the structure of a cell complex in the sense of Whitehead (with a little more trouble the

Received September 30, 1957.

corresponding global result could be obtained, but it is not needed here). Finally, the following three questions of Nash **(1)** are answered. Are sheets closed sets? Does a real algebraic variety have just a finite number of sheets? Does each sheet have a point on it with a neighbourhood containing no points of the variety not on that sheet? The answers are respectively yes, yes, no.

In the local cell decomposition of a variety mentioned above, each cell is contained in a sheet. It is natural to ask when cells with common frontier points belong to the same sheet. The answer is not hard to see when two r-cells meet along a variety of dimension $r - 1$, but the general case seems a bit more difficult, and so far no satisfactory answer has been worked out.

I should like here to draw attention to a recent paper by Whitney **(4)** in which some further connectivity properties of real algebraic varieties are obtained.

PART ONE: PLANE CURVES

1. Definitions. All the curves to be discussed in this paper will be contained in Euclidean spaces; thus, a curve is specified by setting the co-ordinates $x_1, x_2, \ldots, x_n$ in the relevant space equal to continuous functions of a real parameter t, which will, in general, be assumed to vary from 0 to 1. The equations $x_i = f_i(t)$ so obtained are the parametric equations of the curve.

An analytic arc in Euclidean n-space is defined to be an arc with parametric equations $x_i = f_i(t)$, where the f_i are real analytic functions of the variable t, $0 \leqslant t \leqslant 1$.

Let C be an analytic arc in n-space and let P be a point of C, with co-ordinates $(x_1', x_2', \ldots, x_n')$, say. Let t_0 be a value of t for which $f_i(t_0) = x_i'$, where $x_i = f_i(t)$ $(i = 1, 2, \ldots, n)$ are the parametric equations of C. Let $P(t_0)$ be the variable point with co-ordinates $(g_1(t), g_2(t), \ldots, g_n(t))$, where the g_i are the expansions of the f_i in powers of $t - t_0$; a similar definition is to be made for all t_0 such that $f_i(t_0) = x_i'$ $(i = 1, 2, \ldots, n)$. If A is the ideal of C in the ring of power series in the $x_i - x_i'$ with real coefficients, then it is known that A has just a finite number of prime components. The points $P(t_0)$, for all possible t_0, are generic zeros for these prime components, and so t_0 has only a finite number of possible values. That is to say, in a neighbourhood of each of its points, C consists of a finite number of irreducible algebroid branches.

Continuing with the notation of the last paragraph, suppose that, in a neighbourhood of P, C consists of exactly one algebroid branch, and let $F_1, F_2, \ldots, F_r$ be a basis for the ideal of power series in the $x_i - x_i'$ vanishing on C around P. Then P is said to be simple on C if the matrix $(\partial F_i/\partial x_j)$ is of rank $n - 1$ at P. If the rank of this matrix is less than $n - 1$, or if C consists near P of more than one branch, then P is said to be a singular point of C.

If at least one of the functions $f_i(t)$ appearing in the parametric equations of C has a non-zero derivative at P, it is not hard to see that P is a simple point of C. Thus, the singular points must be among those at which the derivatives of all the f_i vanish. Since the f_i are analytic, it is clear that there can be only a finite number of such points. Thus, an analytic arc has at most a finite number of singularities. Note, incidentally, that the derivatives of all the f_i may vanish at a simple point; consider, for example, the origin on the plane curve given by $x = t^2$, $y = t^6$.

It will be assumed for convenience in what follows that the parameter values $t = 0$ and $t = 1$ are always mapped on simple points of any analytic arc under consideration. These points will be called the end-points of the arc.

An algebraic arc will be defined to be an analytic arc which lies entirely on some real algebraic curve.

Let $P_1, P_2, \dots, P_m$ be a finite collection of points of Euclidean n-space and for each pair P_i, P_{i+1} let C_i be an analytic arc with these points as end-points, not passing through any other of the P_j. Then the point-set union C of the C_i will be called a piecewise analytic curve. The C_i will be called the arcs belonging to C and the P_j will be called the joints of C.

A piecewise algebraic curve is a piecewise analytic curve all of whose arcs are algebraic arcs. A variation of this definition is obtained by taking $P_1 = P_n$, when the resulting piecewise analytic or algebraic curve will be called closed. To avoid repetitive descriptions later it will be convenient to agree that the term "singularities of a piecewise analytic or algebraic curve" means the set of singularities of the individual arcs along with all intersections of the arcs other than the joints; thus, the joints of the curve are not counted among the singularities.

Let C and C' be piecewise analytic curves. Then C' will be called an ϵ-approximation of C, for a given positive number ϵ, if there is a homeomorphism $f{:}C \to C'$ such that the distance of $f(p)$ from p is less than ϵ for all $p \in C$.

Let C and C' be piecewise analytic curves and let $p \in C$, $p' \in C'$. Then C and C' are said to be analytically equivalent at the pair p, p' if there are neighbourhoods U and U' of p and p', respectively, and an analytic homeomorphism f of U onto U' which carries $U \cap C$ onto $U' \cap C'$. An analytic homeomorphism will map the point $(x_1, x_2, \dots, x_n)$ on the point $(X_1, X_2, \dots, X_n)$ given by the formulae

$$X_i - a'_i = \sum_{j=1}^{n} a_{ij}(x_j - a_j) + F_i, \qquad i = 1, 2, \dots, n$$

where $(a_1, a_2, \dots, a_n)$ and $(a_1', a_2', \dots, a_n')$ are the co-ordinates of p and p' respectively, the determinant $|a_{ij}|$ is not zero and the F_i are power series in the $x_j - a_j$ of order not less than two (the order of a power series being the degree of the lowest terms appearing). The analytic equivalence will be said to be of order r if f can be so chosen that all the series F_i are of order not less than r.

Let C' be an ϵ-approximation of the piecewise analytic curve C, and let f be the appropriate homeomorphism of C onto C'. Then, if $f(p)$ is simple on

C' whenever p is simple on C, and if $f(p) = p$ for each singular point p of C, and if C and C' are analytically equivalent at each pair $p, f(p) = p$, for each singularity p of C, then C' will be called a singularity preserving ϵ-approximation of C.

The main idea to be treated in what follows is that of the smoothing approximation. This notion will now be introduced in two forms. In the first place let C denote the figure in the (x, y)-plane consisting of two analytic arcs with a common point P simple on each of them, and not an end point of either of them, and assume that the tangents to the two arcs at P are distinct. The curve C' will be called an ϵ-approximation of C smoothed at P if:

(1) There is a continuous mapping $f\colon C' \to C$ such that $f^{-1}(P)$ consists of two distinct points P_1 and P_2 on C', and f is a homeomorphism on $C - P_1 - P_2$;

(2) The distance of p from $f(p)$ is less than ϵ for all $p \in C$;

(3) There is a neighbourhood U of P in the plane and an analytic homeomorphism F of U on a circle V of centre $(0, 0)$ in the (X, Y)-plane such that $F(U \cap C)$ consists of the parts of the X and Y axes in V and $F(U \cap C')$ consists of the part of the hyperbola $XY = 1$ contained in V.

The second form in which this idea will be wanted is as follows. Let C be a piecewise analytic curve in the plane consisting of the union of two analytic arcs C_1 and C_2 with the joint P, and suppose that the tangents to C_1 and C_2 at P are distinct. Then an ϵ-approximation of C smoothed at P is defined as above, with the modification that $F(U \cap C)$ consists of the positive parts of the X- and Y-axes in V, and $F(U \cap C')$ consists of the part of $XY = 1$ in the first quadrant of V.

Smoothing approximations will later be required not only in the plane but also in spaces of any dimension. Let C be a figure in n-space consisting either of two analytic arcs crossing at P or of a piecewise analytic curve with the joint P, the tangents at P being distinct in each case. Then C' will be called an approximation of C smoothed at P if there is an analytic homeomorphism F of a neighbourhood U of P onto a sphere V of centre P' such that $F(U \cap C)$ and $F(U \cap C')$ lie in a plane through P', and $F(U \cap C')$ is an approximation of $F(U \cap C)$, in the sense already defined, smoothed at P'.

2. Analytic equivalence and smoothing. The object of this section is to show how singularity preserving and smoothing approximations of plane curves can be explicitly constructed.

LEMMA 2.1. *Let F be a power series in x and y free of multiple factors, let G be a power series in x and y and let λ be a real number. Assume F and G to be of order* $\geqslant 1$, *so that* $F = 0$ *and* $F + \lambda G = 0$ *are the equations, in a neighbourhood of the origin, of curves C and C', each consisting of a finite number of analytic arcs through the origin. Then:*

(1) *If the integer r is pre-assigned, there is an integer s such that if G is of order $\geqslant s$, C and C' are analytically equivalent at the origin (a self-corresponding point), the analytic equivalence being of order $\geqslant r$.*

(2) *If the analytic equivalence of* (1) *is induced by an analytic homeomorphism $f\colon U \rightarrow U'$, where U and U' are neighbourhoods of the origin, then f depends analytically on λ.*

Proof. The proof of this lemma is due to Samuel **(2)**, with some minor changes. Write

$$F_1 = \frac{\partial F}{\partial x}, \quad F_2 = \frac{\partial F}{\partial y}.$$

In the ring of power series in x and y with real coefficients let $\mathfrak{a}$ be the ideal (F_1, F_2) and let $\mathfrak{p}$ be the ideal (x, y). The ideal $(F, \mathfrak{a})$ has an isolated zero at the origin and so there is an integer d such that $\mathfrak{p}^d \subset (F, \mathfrak{a})$. It follows at once that, for any integer k, $\mathfrak{p}^{2d+k} \subset F\mathfrak{p}^k + \mathfrak{a}^2\mathfrak{p}^k$. Then if G is of order $2d + k$, that is to say, if $G \in \mathfrak{p}^{2d+k}$, $F + \lambda G$ can be written as $F + \lambda H + \lambda FK$, where $H \in \mathfrak{a}^2\mathfrak{p}^k$ and $K \in \mathfrak{p}^k$. $1 + \lambda K$ has an inverse in the ring of power series in x and y, convergent in a neighbourhood of the origin, and so the equation $F + \lambda G = 0$ is equivalent to the equation $F + \lambda H(1 + \lambda K)^{-1} = 0$. The last equation can be written as $F + \Sigma A_{ij}F_iF_j = 0$, where the A_{ij} are in $\mathfrak{p}^k$ and have co-efficients analytic in λ, and vanishing for $\lambda = 0$. It must now be shown that there is an automorphism S of the power series ring in x and y given by equations of the form:

$$S(x) = x + u_{11}F_1 + u_{12}F_2,$$
$$S(y) = y + u_{21}F_1 + u_{22}F_2,$$

where the u_{ij} are power series in x and y vanishing at the origin, such that $S(F) = F + \Sigma A_{ij}F_iF_j$. The existence of S will establish the required analytic equivalence between C and C'.

Now, by Taylor's theorem, $S(F) = F(S(x), S(y)) = F + \Sigma u_{ij}F_iF_j +$ terms of degree $\geqslant 2$ in the u_{ij} and in the F_i. Thus, to prove the existence of S, the u_{ij} must be determined so that

$$\sum u_{ij}F_iF_j + \ldots = \sum A_{ij}F_iF_j,$$

where the dots denote the higher terms. A solution can be obtained by picking out from this equation the terms in F_iF_j for each pair i, j. The following four equations are thus obtained:

$$u_{ij} = A_{ij} + \text{terms of degree} \geqslant 2 \text{ in the } u\text{'s}.$$

These equations can be solved formally by iteration for the u_{ij} as power series in x and y; convergence is assured by the implicit function theorem, provided x and y are small enough.

To see that the analytic equivalence between C and C' obtained in this way satisfies (1) and (2) in the statement of the lemma, note that the orders of the A_{ij} are all $> k$, and so the orders of the u_{ij} also satisfy this inequality. Condition (1) of the lemma follows at once. To verify condition (2), note that the presence of λ in the terms A_{ij} (remembering that the A_{ij} have coefficients analytic in λ vanishing for $\lambda = 0$) implies that the coefficients of the u_{ij}, written as power series in x and y, will be analytic in λ, as required.

It is an immediate corollary of this lemma that, if the neighbourhoods U and U' are fixed so that the equations of the automorphism S, or what is essentially the same, of the homeomorphism $f\colon U \to U'$ are convergent, then, if λ is taken small enough, $C \cap U$ will be approximated arbitrarily closely by $C' \cap U'$.

Lemma 2.2. *Let C be an algebraic curve in the plane with a singular point P at which exactly two simple branches meet with distinct tangents, and let U be a neighbourhood of P such that $U \cap C$ is homeomorphic to two crossed line segments. Let $F = 0$ be the irreducible equation of C and let G be any polynomial in x and y not vanishing at P. Then if ϵ is pre-assigned and λ is taken small enough $F + \lambda G = 0$ is a smoothed ϵ-approximation of C within a sufficiently small neighbourhood U_0 of P.*

Proof. Take P as origin, and let U_0 be a neighbourhood of P such that $G \neq 0$ in U_0. If U_0 is small enough, G^{-1} is a convergent power series in x, y in U_0. Also, if U_0 is small enough, F can be factorized into fgh, where f, g, h are convergent power series in U_0, $f = 0$ and $g = 0$ are the equations of the two branches meeting at P, and $h \neq 0$ at P. f and g are thus of order one. Then the equations $X = f$, $Y = ghG^{-1}$ define an analytic homeomorphism of U_0 onto a neighbourhood of the origin in the (X, Y)-plane. It is not hard to see that if λ is small enough, this homeomorphism defines the required smoothing approximation.

The above lemmas will be combined to give a proof of the following theorem:

Theorem 1. *Let C be a plane algebraic curve with singularities at (x_i, y_i) $(i = 1, 2, \ldots, n)$. Let exactly two simple branches of C with distinct tangents meet at (x_1, y_1). Then, if K is a circular disc containing the (x_i, y_i) and ϵ is a pre-assigned positive number, there exist algebraic curves C_1 and C_2 such that $C_1 \cap K$ and $C_2 \cap K$ are ϵ-approximations of $C \cap K$, smoothed in the two complementary ways at (x_1, y_1) (corresponding to the two complementary hyperbolas $xy = \pm 1$) and otherwise singularity preserving, with analytic equivalence of pre-assigned order at each singularity.*

Proof. Let $F(x, y) = 0$ be the equation of C, free from multiple factors and define the polynomial $G(x, y)$ by the equation

$$G(x, y) = \prod_{i=2}^{n} [(x - x_i)^2 + (y - y_i)^2]^r.$$

G vanishes at each of the (x_i, y_i) for $i = 2, \ldots, n$ and is of order $2r$ at each of these points. And so, by Lemma 2.1, if r is large enough, the curve

$F+\lambda G=0$ is a singularity preserving approximation of $F=0$ in suitably chosen neighbourhoods of the (x_i, y_i) $(i \neq 1)$, and the approximation can be made arbitrarily good by taking λ small enough. Also $G \neq 0$ at (x_1, y_1) and so, by Lemma 2.2, $F+\lambda G=0$ is, in a suitable neighbourhood of (x_1, y_1), a smoothed approximation of $F=0$, and again the approximation can be made arbitrarily good by making λ sufficiently small.

On the other hand, if p is a simple point of C, there is a neighbourhood of p in which there is an admissible system of co-ordinates (in the sense of the real analytic structure of the (x, y)-plane) one of which is the arc length along C while the other is λ. Thus, if γ is a non-singular arc on C, γ has a neighbourhood in which the only part of $F+\lambda G=0$, for λ sufficiently small, is a non-singular arc whose points are at arbitrarily small distance from γ.

Let C' be the curve $F+\lambda G=0$. Then, to complete the proof of the theorem, a mapping f: $C' \to C$ must be constructed to satisfy the conditions of the definitions of smoothing and singularity preserving approximations. In order to construct this mapping take an open covering of $C \cap K$ as follows.

(i) About each singularity (x_i, y_i) take a neighbourhood U_i such that in U_i, for λ sufficiently small, $F+\lambda G=0$ gives a singularity preserving or smoothing approximation of $F=0$, as the case may be. Writing g for the inverse of the mapping f which is to be constructed (and remembering that g will be a mapping, and in fact a homeomorphism on C with (x_1, y_1) removed, the latter point being carried by g into two distinct points), this means that g is now constructed on the $U_i \cap C$.

(ii) $C \cap K$ with the (x_i, y_i) removed consists of a finite number of simple arcs joining singular points to each other, or joining singular points to frontier points of K, or joining frontier points of K to each other. Let γ be one of these arcs and let the end-point p_1 be one of the (x_i, y_i). Let q_1 be the point on γ furthest from p_1 for which the operation g is already defined and let q_1' be a point between p_1 and q_1 on γ. Proceed similarly at the other end p_2 of γ, marking points q_2 and q_2' on γ. If an end point of γ, say p_1, is non-singular on C but lies on the frontier of K, take $q_1=p_1$ and take q_1' near q_1 but outside K. Let $U(\gamma)$ be the union of normal line segments to γ at all points from q_1' to q_2' of length $\delta(\gamma)$ on either side of γ. If $\delta(\gamma)$ is small enough, $U(\gamma)$ is fibred by these normal segments. Repeat this procedure for each arc of $C \cap K$ with the singularities removed.

The U_i along with the $U(\gamma)$ form the required covering of $C \cap K$. Let U be the union of these sets; then U is a neighbourhood of $C \cap K$. Since $F \neq 0$ in $K-U$, it is clear that, for λ sufficiently small, $C' \cap K$ lies entirely in U. Also, fixing attention on the arc γ, and using the notation introduced above, it follows at once from Lemma 2.1 or 2.2, whichever is relevant, that, for λ small enough, $g(q_1)$ and $g(q_2)$ are in $U(\gamma)$. A similar statement can be made for all the other arcs of $C \cap K$. Then, remembering that λ can be taken as one of the local co-ordinates in the plane in $U(\gamma)$, it follows at once that $C' \cap U(\gamma)$ consists of a simple arc with $g(q_1)$ and $g(q_2)$ near its end-points.

It is now easy to see that g can be extended to the whole of $C \cap K$ by mapping the arc q_1q_2 of C on the arc $g(q_1)g(q_2)$ of C' in $U(\gamma)$, proceeding similarly for each arc γ of $C \cap K$. For ϵ pre-assigned, $f = g^{-1}$ is an ϵ-approximation if λ is small enough, and the theorem is completely proved.

The curves C_1 and C_2 mentioned in the statement of the theorem refer to the two complementary ways of smoothing at (x_1, y_1), corresponding respectively to positive and negative values of λ.

It is clear that the above theorem could be modified to yield an approximation of C which is smoothed at several singularities, while preserving the rest.

Also, it is clear that the order of the analytic equivalence between C and C' at each singularity can be made arbitrarily high by taking the exponent r in G large enough.

3. Preliminary approximation theorems. The principal object of this part of the paper is the application of Theorem 1 to the approximation of piecewise analytic curves by circuits of algebraic curves. The way in which this is to be done will now be sketched, the details being completed later in this section and in the next.

Fix attention first on a closed piecewise algebraic curve C in the plane, and let the arcs of C be C_i $(i = 1, 2, \ldots, n)$. Each C_i is part of an algebraic curve $\bar{C}_i$, and C is part of the composite algebraic curve $\bar{C} = \bigcup \bar{C}_i$. If matters are suitably arranged, Theorem 1 can be applied to approximate $\bar{C}$ by an algebraic curve $\bar{C}'$, smoothing at the joins of the C_i. Thus, C is approximated by a circuit C' of the algebraic curve $\bar{C}'$. Parts of $\bar{C}' - C'$ may, however, meet C'. The next step is to show that the curve $\bar{C}'$ can be modified in such a way that the approximating circuit becomes isolated, that is to say, does not meet any other part of the curve. The lemmas which establish the procedure for isolating the approximating circuit will be dealt with first in this section. The sequence of operations can be summarized as follows:

(1) Lift $\bar{C}'$ into 3-space in such a way that C' is separated from the rest.

(2) Make a transformation of 3-space so that the unwanted circuits are removed.

(3) Project back onto the plane.

LEMMA 3.1. *Let C be a real plane algebraic curve and let (x_i, y_i), $(i = 1, 2, \ldots, r)$, be singularities at each of which exactly two simple branches meet with distinct tangents (C may, of course, have other singularities as well). Then there is a curve C' in 3-space which, under the projection P onto the (x, y)-plane, projects onto C and is such that $P^{-1}(x_i, y_i)$, for each i, consists of exactly two distinct points, while, apart from these points, P is a homeomorphism on C'. Also, if (x_0, y_0) is a singularity of C other than the (x_i, y_i) and if $P^{-1}(x_0, y_0)$ $= (x_0, y_0, z_0)$, then C' and C are analytically equivalent at the pair of points (x_0, y_0, z_0) and $(x_0, y_0, 0)$, to an arbitrarily high order.*

Proof. Assume co-ordinates chosen so that the lines $x = x_1$, $x = x_2, \ldots,$ $x = x_r$ meet C in simple points, apart from the points (x_i, y_i) $(i = 1,2,3, \ldots, r)$ and suppose that at all these simple points the tangents are not parallel to the y-axis. Also, at each of the points (x_i, y_i) $(i = 1, 2, \ldots, r)$ there are two tangents, and it is to be assumed that none of these tangents is parallel to the y-axis. Let Y_1, $Y_2, \ldots, Y_s$ be the y-co-ordinates of the singularities of C, other than the (x_i, y_i). Let the line $x = x_i$ meet C at the simple points (x_i, y_{ij}), $j = 1, 2, \ldots, m$, and let f be the product of all the distinct expressions picked from the $y - Y_i$, the $y - y_i$ and the $y - y_{ij}$, for all i, j; that is to say, if a number of these expressions should happen to be equal, the corresponding factor of f is nevertheless to appear just once. Similarly, let g be the product of distinct factors picked from the set $x - x_i$. Let $F(x, y)$ be the rational function f/g.

Examining the behaviour of F on the curve C, it is clear that the only possible points of indeterminacy are the zeros of g, namely, the points (x_i, y_i) and (x_i, y_{ij}) for all i, j. Representing y as a power series in $x - x_i$ for points of C around (x_i, y_{ij}) it turns out that F is continuous on C at that point. On the other hand, at the point (x_i, y_i) there are two distinct branches of C with distinct tangents. Making use of the two corresponding expansions of y in powers of $x - x_i$, it follows this time that F is continuous on each of the branches of C at (x_i, y_i) taken separately. The fact that the tangents to these two branches are distinct ensures that the two limits of F as (x, y) approaches (x_i, y_i) along the two branches of C are different.

Now let C' be the curve in 3-space whose points are of the form

$$(x, y, F^m(x, y)),$$

where (x, y) is on C. Then C' is the curve required by the statement of the present lemma. For if P is the projection of C' on C it is clear that $P^{-1}(x_i, y_i)$ consists of two distinct points, corresponding to the two limits of F along the branches at (x_i, y_i) and that P is one-one on C' except at these points which project doubly. To complete the proof of the lemma, the behaviour of P at points projecting on singularities of C other than the (x_i, y_i) must be examined. Around such a singularity (x_0, y_0), F is a real analytic function of x and y, and so P extends to an analytic homeomorphism of a neighbourhood of

$$(x_0, y_0, z_0) = P^{-1}(x_0, y_0)$$

on a neighbourhood of $(x_0, y_0, 0)$ in 3-space. For example, the mapping of (x, y, z) on $(x, y, z - F^m(x, y))$ gives such an extension which is an analytic equivalence of order m. This completes the proof.

LEMMA 3.2. *Let C be an algebraic arc in n-space E_n with a singularity Q projecting, under rectangular projection, on an arc C' in r-space E_r with a singularity at Q'. Let C and C' be analytically equivalent at Q, Q', the equivalence being induced by an analytic homeomorphism F of a neighbourhood U of Q on a neigh-*

bourhood U' of Q', such that, on $F^{-1}(U' \cap E_r)$, F coincides with the projection. Then a sufficiently good approximation C_1 of C projects on an arbitrarily good approximation C_1' of C'. Also, if C_1 is a singularity preserving approximation of C with analytic equivalence of sufficiently high order, and if the order of the analytic equivalence induced by F is of sufficiently high order, then C_1' will be a singularity preserving approximation of C', and the corresponding analytic equivalence at Q' will be of pre-assigned order.

Proof. The first part of the lemma, that a sufficiently good approximation of C projects into an arbitrarily good approximation of C', is practically trivial, and so attention will be fixed on the second part. Let P be the orthogonal projection of E_n on E_r. Let F_1 be an analytic homeomorphism of U (the same neighbourhood as in the above statement; no generality is lost as U can, if necessary, be shrunk to suit both situations) on a neighbourhood U_1 of Q, such that F_1 induces the analytic equivalence assumed between C and C_1 at Q. Let U_0 be a neighbourhood of Q' in E_r, say $U' \cap E_r$. Define F' as $PF_1 F^{-1}$ restricted to U_0. It is clear that F' defines an analytic equivalence of C_1' and C' at Q' as required. Also, it is not hard to see that, since P is a linear transformation, the order of F' can be made as high as one pleases by making those of F_1 and F large enough.

LEMMA 3.3. *Let $P_1, P_2, \dots, P_r$ be a set of points in n-space E_n contained in a sphere A with centre the origin. Let P be a point outside A and let U be a pre-assigned neighbourhood of P. Then there exists a rational mapping F of E_n onto itself such that:*

(1) *F approximates the identity mapping arbitrarily closely on A;*

(2) *F carries all points outside a sufficiently large sphere B into U;*

(3) *$F(P_i) = P_i$ for each i, and if $(a_{i1}, a_{i2}, \dots, a_{in})$ are the co-ordinates of P_i then the equations of F in a neighbourhood of P_i are of the form $X_j = x_j + F_{ij}$, where F_{ij} is a power series in the $x_j - a_{ij}$ of pre-assigned order.*

Proof. The idea involved here is similar to that of (**3**, Lemma 2, §3). The mapping constructed there is the composition of stereographic projection of E_n on a sphere in E_{n+1} and an oblique projection back onto E_n. Such a mapping would be rational and would satisfy (1) and (2) above, but not (3). The required mapping will be constructed by making a suitable modification of stereographic projection.

Let r be an integer such that $2r$ is greater than the pre-assigned orders referred to in (3) of the present theorem, let $d(P_i, x)$ be the distance of the point $(x_1, x_2, \dots, x_n)$ from P_i, and define the polynomial

$$G(x) = \prod_i d^{2r}(P_i, x).$$

Then the equations of the mapping F are to be:

$$(1) \qquad \begin{aligned} X_1 &= (k^2x_1 + kG(x)\tan\alpha)/(k^2 + G(x)) \\ X_i &= k^2x_i/(k^2 + G(x)), \qquad i = 2, 3, \dots, n, \end{aligned}$$

where $(X_1, X_2, \ldots, X_n)$ is the transform of $(x_1, x_2, \ldots, x_n)$ under F. It must be shown that the constants k and α can be chosen so that the conditions of the lemma are satisfied.

In order to do this, it is convenient to consider the geometrical meaning of the mapping F with the equations (1). Let the n-space E_n be the hyperplane $x_{n+1} = 0$ in $(n+1)$-space, and let the co-ordinates be chosen so that P is on the x_1-axis. Let α be the angle between the x_{n+1}-axis and the line joining P to the point $(0, 0, \ldots, 0, k)$ in $(n+1)$-space. Then F is the composition of the mappings f and g where $f(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_{n+1})$ and $g(y_1, y_2, \ldots, y_{n+1}) = (X_1, X_2, \ldots, X_n)$ these mappings being defined by the equations:

$$\begin{aligned} &(2) & y_i &= k^2x_i/(k^2 + G(x)), & i &= 1, 2, \ldots, n, \\ & & y_{n+1} &= kG(x)/(k^2 + G(x)). \end{aligned}$$

$$\begin{aligned} &(3) & X_1 &= y_1 + y_{n+1} \tan \alpha \\ & & X_i &= y_i, & i &= 2, \ldots, n. \end{aligned}$$

The mapping g is, of course, the projection of E_{n+1} onto E_n along lines parallel to the join of P and $(0, 0, \ldots, 0, k)$ while f, on the other hand, is a mapping similar to stereographic projection, and would coincide with it if G were replaced by $\sum x_i^2$. f can be described geometrically as the projection from $(0, 0, \ldots, 0, k)$ of E_n on the hypersurface H in E_{n+1} having the equations (2) as parametric equations. f is a one-one mapping of E_{n+1} on H, and in fact a birational correspondence, the inverse mapping being given by

$$x_i = ky_i/(k - y_{n+1}), \qquad i = 1, 2, \ldots, n.$$

Comparing H with the sphere of centre $(0, 0, \ldots, 0, \frac{1}{2}k)$ and radius k in E_{n+1}, given by the parametric equations

$$\begin{aligned} Y_i &= k^2x_i/(k^2 + \sum x_j^2), \qquad i = 1, 2, \ldots, n, \\ Y_{n+1} &= k \sum x_j^2/(k^2 + \sum x_j^2), \end{aligned}$$

it is easy to see that, for points $(x_1, x_2, \ldots, x_n)$ in some bounded set in E_n the distance of $(y_1, y_2, \ldots, y_{n+1})$ from $(Y_1, Y_2, \ldots, Y_{n+1})$ can be made as small as one likes by taking k large enough. Also it is a simple computation to show that, still confining $(x_1, x_2, \ldots, x_n)$ within a bounded set the partial derivatives of y_{n+1} with respect to the x_i, calculated from the equations (2), are as small as one pleases if k is sufficiently large, while $\partial y_j/\partial x_i$ approximates δ_{ij}, the Kronecker δ, for large k. It follows that the normal to H at points corresponding to values of $(x_1, x_2, \ldots, x_n)$ in a bounded set will make an arbitrarily small angle with the x_{n+1}-axis in E_{n+1} if k is large enough. Taking the sphere A as the bounded set in these remarks, it follows at once that a line parallel to the join of P to $(0, 0, \ldots, 0, k)$ will, if k is large enough, meet H at not more than one point corresponding to values of $(x_1, x_2, \ldots, x_n)$ in A. That is to say, $g \circ f = F$ is one-one on A for sufficiently large values of k.

An easy calculation shows that for k sufficiently large, F also approximates the identity mapping arbitrarily closely on A. Thus (1) in the statement of the lemma is proved. To verify (2) note that outside a large sphere B, the polynomial $G(x)$ will be large, and so, if B is large enough, the distance of $(y_1, y_2, \ldots, y_{n+1})$, given by equations (2), from $(0, 0, \ldots, 0, k)$ will be less than a pre-assigned number. On the other hand, a sufficiently small neighbourhood of $(0, 0, \ldots 0, k)$ will project under g into U, and so (2) is proved. To prove part (3) of the lemma, note that $G(x)$, as a power series in the $x_j - a_{ij}$ around P_i, is of order $2r$, and so the series expansions of the X_j in equations (1) are of the required form.

LEMMA 3.4. *Let C be a real algebraic curve and let $C = C_1 \cup C_2$ where C_1 is a closed circuit and C_2 is the rest of the curve. Assume that the intersections of C_1 and C_2 are all double points at which exactly two simple branches meet. Then there exists an algebraic curve $C' = C_1' \cup C_2'$ of which the circuit C_1' is a singularity preserving ϵ-approximation of C_1 for pre-assigned ϵ, while C_2' is contained in an arbitrary neighbourhood of a pre-assigned point P, arbitrarily far from C_1. In addition, the order of the analytic equivalence of C_1 and C_1' at each singularity can be made greater than a pre-assigned integer.*

Proof. The first step is to apply Lemma 3.1 to C, taking the points of intersection of C_1 and C_2 as the (x_i, y_i). In this way a space curve $K = K_1 \cup K_2$ is obtained such that $K_1 \cap K_2 = \phi$. Also, K_1 projects in a one-one manner on C_1, and if (x_0, y_0, z_0) on K_1 projects on a singularity of C_1, then K_1 and C_1, regarded as space curves, are analytically equivalent to an arbitrarily high order at the pair (x_0, y_0, z_0) and $(x_0, y_0, 0)$. The next step is to move K_2 to a great distance from K_1. Take the 3-space with co-ordinates (x, y, z) as the hyperplane $t = 0$ in (x, y, z, t)-space. Let A and B be spheres in the latter space with the origin as centre, and such that $K_1 \subset A \subset B$. Let B be of radius R. Let $\phi(x, y, z)$ be a continuous function equal to zero on K_1 and equal to $2R$ on $K_2 \cap B$. Apply the Weierstrass approximation theorem to approximate ϕ on B by means of a polynomial f. Let g be a rational function of x, y, z which is equal to zero to a pre-assigned order at each singularity of K_1, satisfies everywhere the inequality $0 < g < 1$, and also satisfies $g > 1 - \delta$ outside pre-assigned neighbourhoods of the singularities of K_1, δ being a pre-assigned positive number. A method of construction for g will be given in Lemma 3.5 below. Now, for each point (x, y, z) of K, set $G(x, y, z)$ equal to the point in 4-space with co-ordinates (x, y, z, fg), where f and g are evaluated at (x, y, z). Then G is a birational transformation of K into a curve

$$K' = K_1' \cup K_2'.$$

K_1' projects in a one-one manner on K_1 and is contained in A, while K_2' lies outside B, if the functions f and g have been suitably chosen. Also, there is analytic equivalence between K_1 and K_1' at each singularity, the order being that of g, regarded as a power series at the point in question.

The mapping F of Lemma 3.3, carrying (x, y, z, t)-space on itself is now to be applied. F carries K' into a curve

$$K'' = K_1'' \cup K_2''.$$

The points P_i of Lemma 3.3 are here to be taken as the singularities of K_1'. Now, the notation used here is intended to indicate that K_1'' and K_2'' are the images of K_1' and K_2' under F; it must, however, be checked that K'' constitutes the whole of a real algebraic curve. That this is so follows at once from the fact that F is birational on K', being the composition of the birational mapping f of Lemma 3.3 and the projection g of Lemma 3.3 which, being one-one on A, is certainly birational on $f(K')$. By Lemma 3.3 K_1'' is a singularity preserving approximation of K_1', which will be arbitrarily close if the constant k of Lemma 3.3 is taken large enough. Also, the analytic equivalence of K_1' and K_1'' at each singularity can be made of arbitrarily high order. If the sphere B has been made large enough, Lemma 3.3 implies that K_2'' will be contained in a preassigned neighbourhood of P in 4-space.

The proof will now be completed by a projection back onto the plane, which is the subset $z = t = 0$ of 4-space, applying Lemma 3.2 to obtain the required result. To apply Lemma 3.2, note first that, under the orthogonal projection of (x, y, z, t)-space on the (x, y)-plane K_1' is carried onto C_1, the singularities of the former being mapped on those of the latter, with analytic equivalence in each case of pre-assigned order. Also, by the above argument, K_1'' is a singularity preserving approximation of K_1', with analytic equivalence at each singularity of arbitrary high order. It follows at once from Lemma 3.2, that, if K_1'' is a sufficiently close approximation of K_1', and if the orders of analytic equivalence just mentioned are high enough, then C_1', the projection of K_1'' is as stated in this lemma. Also, K_2'' is contained in a neighbourhood of P, and so the same holds for its projection C_2', and the proof of the lemma is complete.

Lemma 3.5. *Let $P_1, P_2, \ldots, P_m$ be points in Euclidean space of any dimension, and let $U_1, U_2, \ldots, U_m$ be spheres with these points as centres and radii $r_1, r_2, \ldots, r_m$, respectively. Then there is a rational function f such that f vanishes to a pre-assigned order at each P_i and $f > 1 - \epsilon$ outside the U_i, for a pre-assigned positive number ϵ, and at all points* $0 < f < 1$.

Proof. Let η be a positive number and let s be a positive integer. Denote by $d(P_i, x)$ the distance of P from the point with co-ordinates $(x_1, x_2, \ldots, x_n)$. Define

$$f_i(x) = d^{2s}(P_i, x)/(\eta r^{2s} + d^{2s}(P_i, x)).$$

Clearly $0 < f_i < 1$, and also f_i vanishes at P_i, being of order s there (regarded as a power series around P_i). It is not hard to verify that $f_i > 1 - \eta$ at points outside U_i. Now set

$$f(x) = \prod_{i=1}^{m} f_i(x).$$

It is easy to see that if η is small enough, f satisfies the requirements of the lemma.

4. The main approximation theorems in the plane. The main results of Part I will now be obtained, namely, approximation theorems for piecewise analytic and algebraic curves in the plane. Attention will first be fixed on closed curves, and the required result approached in two stages, namely, Theorems 2 and 3.

THEOREM 2. *Let C be a closed piecewise algebraic curve with arcs C_i, joints P_j, satisfying the following conditions:*

(1) *C_i is part of an algebraic curve $\bar{C}_i$ and at each P_i exactly two C_j meet, namely C_i, and C_{i+1}, and P_i is to be simple on both, the two tangents being distinct.*

(2) *If $\bar{C}_i - C_i$ meets C_j (this is to include the case $i = j$) then it does so at a point which is simple both on C_j and on $\bar{C}_i - C_i$ and the two tangents there are distinct.*

Then there exists an algebraic curve $C' = C_1' \cup C_2'$, where the circuit C_1' is an ϵ-approximation of C (ϵ being pre-assigned) smoothed at the P_i and singularity preserving, with analytic equivalence of pre-assigned order at the singularities, while C_2' is contained in an arbitrarily pre-assigned set U.

Proof. Apply Theorem 1 to $\bar{C} = \bigcup \bar{C}_i$. Within a disc containing C this gives an arbitrarily good approximation of C, smoothing at the P_i, and otherwise singularity preserving with analytic equivalence of arbitrarily high order at the singularities. Let the resulting curve be $C^* = C^*_1 \cup C^*_2$ where C^*_1 approximates C. Then the above conditions ensure that the intersections of C^*_1 and C^*_2 are all points where two simple branches meet with distinct tangents. Applying Lemma 3.4 to C^*, the required result follows.

Theorem 3 will now generalize Theorem 2 and remove restrictions placed temporarily on the C in that Theorem.

THEOREM 3. *Let C be a closed piecewise analytic curve in the plane with arcs C_i, and joints P_j. Then there exists an algebraic curve $C_1' \cup C_2'$ with one circuit C_1' giving an ϵ-approximation of C, where ϵ is pre-assigned, smoothed at the P_i, and otherwise singularity preserving with analytic equivalence of arbitrarily high order, while C_2' lies in a pre-assigned set.*

Proof. Let Q_i be a singular point of C; it may be a singular point of just one C_j or a point at which several of these arcs meet, being either singular or simple on each one of these. Let U_i be a sufficiently small neighbourhood of Q_i and let $F_i = 0$ be the equation of $C \cap U_i$, where F_i is a power series in x and y. If all sufficiently high powers of x and y in F_i are dropped a polynomial equation $F_i' = 0$ is óbtained which represents in U_i a curve analytically equivalent to $C \cap U_i$, the analytical equivalence being of arbitrarily high order. The same can be said of $F_i' + \lambda_i G_i = 0$ where G_i is a polynomial containing only sufficiently high powers of x and y and λ_i is a real number.

Repeat the above procedure at each singularity Q_i of C, thus obtaining a set of algebraic curves $F_i' + \lambda_i G_i = 0$. By suitable adjustment of the λ_i it may be ensured that the curve $F_i' + \lambda_i G_i = 0$ does not contain Q_j for $i \neq j$. Now approximate the remainder of C outside the neighbourhoods U_i by straight line segments. These segments along with the parts of the curves $F_i' + \lambda_i G_i = 0$ in U_i are to play the part of the C_i of Theorem 2. It is easy to see that these line segments can be chosen in such a way that the conditions of that theorem hold. An application of Theorem 2 then gives the required result.

COROLLARY. *Theorem* 3 *will also hold if C is replaced by any piecewise analytic curve, not necessarily closed.*

Proof. For if the C_i are the arcs of a non-closed piecewise analytic curve then a closed curve may be obtained by joining the first and last end points of C by any analytic arc not meeting the C_i at any other point. Theorem 3 may then be applied to the resulting closed curve after which the unwanted portion of the approximation, corresponding to the additional analytic arc, can be discarded.

PART TWO: CURVES ON REAL ALGEBRAIC VARIETIES

5. Approximation of a curve in 3-space. It has now been shown that a piecewise analytic curve in the plane, that is to say, a sequence of analytic arcs joined end to end, can be approximated by part of an algebraic curve. As indicated in the introduction, that approximation theorem was the first stage towards a similar approximation theorem on a real algebraic variety. The second stage is to generalize the result in the plane to Euclidean spaces of higher dimension. The general result will be proved by induction on the dimension of the surrounding space starting with dimension 3. The case of dimension 3 needs special attention because it is not, in general, possible to make a projection of a curve on to a plane in a one-one manner.

The procedure for obtaining the required approximation in 3-space is as follows.

(1) Let C be a closed piecewise analytic curve in 3-space. Let C_y and C_z be its projections on the planes $y = 0$ and $z = 0$ respectively. Approximate these curves by circuits of algebraic curves C_y', C_z' with equations $f(x, z) = 0$, $g(x, y) = 0$, respectively. These approximations are to be smoothed at the joints and otherwise singularity preserving. Then it will turn out that in 3-space the curve C' with equations $f = g = 0$ has a circuit approximating C arbitrarily closely, smoothing at the joints and otherwise singularity preserving (provided that the approximations in the two planes are suitably made).

(2) Now C' can be represented alternatively by the equations $g(x, y) = 0$, $z = h(x, y)$ where h is a rational function with indeterminacies at points on which more than one point of C' or some singularity of C' projects. In particular the circuit of C' approximating C is obtained by allowing the argument

(x, y) of h to vary on a suitable circuit of C_z'. Approximate this circuit of C_z' by an isolated circuit of an algebraic curve C'' with equation $g'(x, y) = 0$ and then it will be shown that the equations $g' = 0, z = h$ define a curve C'' in space with an isolated circuit approximating C, smoothing at the joints and otherwise singularity preserving.

The details of the procedure sketched in (1) above will be carried out by a sequence of lemmas in the next three sections. These treat in turn various special points of C. First, there are the singularities of C which are to be preserved; second, points at which C is to be smoothed; and, finally, points at which the tangent to C is parallel to the (x, z) - plane. An example of the last type is the point $(0, 0, 0)$ on one loop of the intersection of the cylinders

$$x^2 + (y - 1)^2 = 1, \qquad z^2 + (y - 1)^2 = 1.$$

A slight displacement of the cylinders, unless subject to suitable restrictions, would clearly change the configuration at the origin.

6. Analytic equivalence.

The object of this section is to give a criterion for the analytic equivalence of curves defined locally by suitably related sets of equations. As the results are to be used later in the proofs of the general approximation theorems, it will be convenient to state a lemma applicable to space of any dimension.

LEMMA 6.1. *Let $f_i(x)$, $i = 1, 2, \ldots, m$, be power series in $x_1, x_2, \ldots, x_n$, and let $F(x, z)$ be a power series in $x_1, x_2, \ldots, x_n, z$; similar meanings are to be attached to $f_i'(x)$, $i = 1, 2, \ldots, m$, and $F'(x, z)$. Suppose that there is an automorphism S of the power series ring in the x_i of the type $S(x_i) = x_i + h_i(x)$, where the h_i are power series of order not less than 2, and where $S(f_i)$ is a linear combination of the f_j', and $S^{-1}(f_i')$ a linear combination of the f_j, for each i. Then, if the orders of the h_i and of the difference $F - F'$ are sufficiently high and if the series $f_1, f_2, \ldots, f_m$, F, and F_z have no common zero in a neighbourhood of the origin, there will be an automorphism T of the power series ring in $x_1, x_2, \ldots, x_n$ and z, of the form*

$$T(x_i) = x_i + h_i(x), T(z) = z + l(x, z),$$

where the order of $l(x, z)$ is greater than a pre-assigned integer, and where T and T^{-1} carry each of the sets $f_1, f_2, \ldots, f_m$, F and $f_1', f_2', \ldots, f_m'$, F' into linear combinations of the other.

Proof. The power series $l(x, z)$ mentioned in the statement of the lemma is to be found in such a way that the set of equations $f_i(x + h) = 0 (i = 1, 2, \ldots, m)$, $F(x + h, z + l) = 0$ is a set of linear combinations of the equations $f_i'(x) = 0 (i = 1, 2, \ldots, m)$, $F'(x, z) = 0$, where $x + h$ denotes the set $x_j + h_j$ $(j = 1, 2, \ldots, n)$. Since the automorphism S already relates the f_i and the f_i' in this way, it will be sufficient to find l in such a way that

$$F(x + h, z + l) = \sum A_i f_i'(x) + BF'(x, z),$$

where the A_i and B are power series in the x_i and z, and B is invertible in a neighbourhood of the origin. The last condition is equivalent to saying that B does not vanish at the origin. Applying Taylor's theorem to the equation just written, it turns out that the power series l, the A_i and B must be determined to satisfy the following equation:

$$(1) \qquad F(x + h, z) + lF_z(x + h, z) + \tfrac{1}{2}l^2 F_{zz}(x + h, z) + \dots$$
$$= \sum A_i f_i'(x) + BF'(x, z)$$

where the dots denote terms involving higher powers of l.

There are two cases to consider. If $F_z \neq 0$ at the origin, then $F_z(x + h, z)$ is invertible around the origin, and so equation (1) can be divided by it. Equation (1) can then be solved by setting the A_i all equal to zero and $B = 1$, and then calculating l iteratively. The first approximation to l will be $F'(x, z) - F(x + h, z)$, and the order of this, which will also be the order of l, can be made arbitrarily high by making the orders of the h_i and $F - F'$ high enough.

The second and more complicated case is where F_z is zero at the origin. In this case, in order to enable an iterative solution for l to be carried out, it will first be shown that the A_i and B in (1) can be chosen in such a way that the terms free from l, namely, $- F(x + h, z) + \sum A_i f_i'(x) + BF'(x, z)$, will be equal to a multiple of $F_z^s(x + h, z)$ for some pre-assigned integer s. By the hypothesis of the operation of the automorphism S on the f_i, this is equivalent to finding power series A_i' and B such that $- F(x + h, z) + \sum A_i' f_i(x + h) + BF'(x, z)$ is a multiple of $F_z^s(x + h, z)$.

In solving this auxiliary problem, it is convenient to make a change of notation, writing X_i for $x_i + h_i$. Thus, the auxiliary problem is as follows: if the $p_i(X)$ $(i = 1, 2, \dots, n)$ are power series of sufficiently high order, then it is required to find power series P_i, Q and R, Q being invertible, such that

$$(2) \qquad - F(X, z) + \sum P_i f_i(X) + QF'(X + p, z) = RF_z^s(X, z).$$

A solution will actually be found in which $Q = 1 + Q''$, where Q'' is of order > 1. Writing $P_i' = - P_i Q^{-1}$, $Q^{-1} = 1 + Q'$ and $R' = Q^{-1}R$, equation (2) becomes

$$(3) \quad F'(X + p, z) = F(X, z) + \sum P_i' f_i(X) + Q'F(X, z) + R'F_z^s(X, z)$$

where this equation is to be solved for the P_i', Q' and R'. Now, since by hypothesis the f_i, F and F_z have an isolated common zero at the origin, there is an integer d such that all monomials of degree d in the X_i and z are in the ideal generated in the ring of power series in the X_i and z by the $f_i(X)$, $F(X, z)$ and $F_z^s(X, z)$. On the other hand, if the p_i are of sufficiently high order, $F'(X + p, z) - F(X, z)$ will be of order $> d$. It follows at once that the P_i', Q' and R' can be chosen so that (3) is satisfied.

Returning to the main question, the solution of this auxiliary problem implies that equation (1) can be rewritten as

$$lF_z(x+h,z) + \tfrac{1}{2}l^2F_{zz}(x+h,z) + \ldots = C(x,z)F_z^s(x+h,z) \tag{4}$$

where C is some power series. This equation will now be solved by setting

$$l = u(x,z)F_z^{s-2}(x+h,z),$$

where u is a power series to be determined. By means of this substitution (4) becomes

$$u + G = CF_z(x+h,z)$$

where G is a power series in the x, z and u, involving only powers greater than the first of u. The power series u can now be obtained from this equation by iteration, the convergence of the process being assured by the implicit function theorem. Note that the first approximation to u in the iterative process is $CF_z(x+h,z)$, and so the first approximation to l is $CF_z^{s-1}(x+h,z)$. Since F_z is zero at the origin, this ensures that l will be of pre-assigned order, merely by taking s big enough. By following through the above proof step by step it is not hard to see the

COROLLARY. *If the h_i and F' depend analytically on one or more parameters in such a way that the h_i vanish and F' reduces to F when these parameters are set equal to zero, then the series l will also depend analytically on these parameters and will vanish when they are set equal to zero.*

Restrict attention now to curve branches in 3-space.

LEMMA 6.2. *Let C be a curve branch (not necessarily irreducible) in a neighbourhood of the origin in 3-space. C is part of a curve with equations $F(y,z) = 0$, $G(x,y) = 0$ where F and G are power series free of double factors. Then, if C' has equations $F'(y,z) = 0, G'(x,y) = 0$ where F' and G' are power series such that $F - F'$, $G - G'$ are of sufficiently high order, it will follow that C and C' are analytically equivalent to an arbitrarily high order at the origin.*

Proof. By Lemma 2.1 there exists a transformation of the type $(x,y) \to (x+h, y+k)$ carrying G into G' where h and k are power series in x and y whose orders are greater than a pre-assigned integer if G and G' are of sufficiently high order. Also, F and F_z have a common isolated zero at the origin. The result then follows from Lemma 6.1.

COROLLARY. *If F' and G' depend analytically on one or more parameters in such a way that they reduce to F and G when these parameters vanish, then the equations of the analytic equivalence referred to in this lemma will also depend analytically on these parameters and will reduce to the identity transformation when the parameters vanish.*

This follows itself from the corollary of Lemma 6.1.

7. Smoothing in 3-space. Let C_1 and C_2 be two analytic arcs in 3-space with common end point P where P is simple on both arcs and projects into simple points of the projections of these arcs on the $(x, y) =$ plane and on the $(y, z) =$ plane. Also, the tangents at P to C_1 and C_2 are to be distinct and are to project into distinct tangents under the above projections. It will be assumed that these tangents are not parallel to the $(x, z) =$ plane. There are two cases to consider.

(1) If P has co-ordinates (x_0, y_0, z_0) then C_1 and C_2 both lie on the same side of the plane $y = y_0$ for a neighbourhood P.

(2) C_1 and C_2 lie on opposite sides of $y = y_0$ around P.

Case (2) can be dealt with straightforwardly, but around points presenting case (1) a modification will be made so that in fact only points presenting case (2) arise.

Suppose case (1) holds. If y_1 is suitably chosen near y_0, $y = y_1$ cuts C_1 and C_2 at uniquely defined points P_1 and P_2 respectively near P. Let C_3 be the circular arc $P_1 P P_2$. It is not hard to see that the arcs C_1 and C_3 at P_1 satisfy all the conditions stated above and present case (2). A similar statement may be made about C_3 and C_2 at P_2. In the sequel all points at which case (1) holds will be dealt with in this way. Attention from now onwards can be confined to case (2).

Let $F_1 = 0$ and $F_2 = 0$ be the irreducible power series equations for the projections C_1', C_2' of C_1, C_2 on the $(x, y) =$ plane and write $F = F_1 F_2$. Let $G(x, y)$ be a power series $\neq 0$ at the projection (x_0, y_0) of P. $C_1' \cup C_2'$ has a smooth approximation of the form $F + \lambda G = 0$, where λ is small, the sign being chosen so that C_1' and C_2' are joined up (cf. Lemma 2.2). Examining this smoothing process again it is required to show that the lines $y = y_1$ meet $F + \lambda G = 0$ in just two points, one of which lies on the smooth approximation of $C_1' \cup C_2'$ where λ is sufficiently small and y_1 is sufficiently near y_0.

To do this, suppose that the linear terms of F_1 and F_2, written in powers of $x - x_0$, $y - y_0$ are

$$a_{11}(x - x_0) + a_{12}(y - y_0)$$
$$a_{21}(x - x_0) + a_{22}(y - y_0).$$

The condition that the tangents to C_1 and C_2 at P should not be parallel to the (x, z)-plane implies that a_{11} and a_{21} are both $\neq 0$. It follows that for small λ, $F + \lambda G = F_1 F_2 + \lambda G$ contains a term in $(x - x_0)^2$. The Weierstrass preparation theorem implies that there is a power series P in $x - x_0$, $y - y_0$ and λ not vanishing at $x = x_0$, $y = y_0$, $\lambda = 0$ such that

$$P(F_1F_2 + \lambda G) \equiv (x - x_0)^2 + a(y)(x - x_0) + b(y) \equiv F'(x, y, \lambda),$$

where a and b are power series in y and λ. The above equation shows that

$$\frac{\partial F'}{\partial \lambda} \neq 0 \text{ at } x - x_0 = y - y_0 = \lambda = 0.$$

And so b contains a linear term in λ, which means that, at $y = y_0$, the discriminant of F', regarded as a quadratic in $x - x_0$, changes sign as λ changes sign. It follows at once that for a suitable choice of sign for λ, $F' = 0$ will have two real roots. It is clear that one of them will lie on the smooth approximation of $C_1' \cup C_2'$. Summing up:

LEMMA 7.1. *If case* (2) *described above holds and if C_1', and C_2', are the projections of C_1, and C_2, respectively, on the (x, y)-plane and (x_0, y_0) is the projection of P, then, in a neighbourhood of (x_0, y_0) there exists an arbitrarily good smooth approximation of $C_1' \cup C_2'$ cut in one point by each line $y = y_1$.*

Suppose that $H = 0$ is the equation of the projection $C_1'' \cup C_2''$ of $C_1 \cup C_2$ on the (y, z)-plane. Then the lines $y = y_1$, $z = z_1$, with (y_1, z_1) near (y_0, z_0) cut the sheet of the surface $F + \lambda G = 0$ over the above constructed approximation of $C_1' \cup C_2'$ in one point. A homeomorphism ϕ of a neighbourhood of $(0, y_0, z_0)$ on the (y, z)-plane and a neighbourhood of (x_0, y_0, z_0) on that surface is thus defined. It is thus clear that for μ small enough and of suitable sign and $K \neq 0$ at (y_0, z_0), K being a power series in $y - y_0$, $z - z_0$, the surface $H + \lambda K = 0$ cuts the (y, z)-plane in a smooth approximation of $C_1'' \cup C_2''$. And so, applying the homeomorphism ϕ, $F + \lambda G = 0$, $H + \mu K = 0$ is a curve an arc of which is a smooth approximation of $C_1 \cup C_2$ in a neighbourhood of P.

The result so obtained may be summed up in

LEMMA 7.2. *Let C_1, C_2 be analytic arcs with the common end point P, case* (2) *holding, and suppose that $C_1 \cup C_2$ is part of the curve $F(x, y) = 0$, $H(y, z) = 0$. Then, in a sufficiently small neighbourhood U of P and for λ, μ sufficiently small and with suitable signs, the curve $F + \lambda G = 0$, $H + \mu K = 0$, where $G(x, y)$ and $H(y, z)$ are analytic functions $\neq 0$ at P, has an arc which is a smooth ϵ-approximation of $(C_1 \cup C_2) \cap U$, where ϵ is pre-assigned.*

8. Tangents parallel to the (x, z)**-plane.** Let P be a simple point of an analytic arc C and suppose that the tangent at P is parallel to the (x, z)-plane. Let the projections of C around P on the (x, y)-plane and (y, z)-plane have respectively the equations $F(x, y) = 0$ and $H(y, z) = 0$. The surfaces $F = 0$ and $H = 0$ touch at P and so the procedure of §7 would not give an approximation of C. The procedure to be followed here is to construct a curve through P analytically equivalent to $F = H = 0$.

In more detail, let (x_0, y_0, z_0) be co-ordinates of P and let G and K be power series in $(x - x_0, y - y_0)$ and $(y - y_0, z - z_0)$, respectively, of sufficiently high order. The conditions of Lemma 6.2 are satisfied by $F = H = 0$; it follows at once from that lemma and its corollary that $F + \lambda G = 0$, $H + \mu K = 0$ is analytically equivalent to $F = H = 0$ in a neighbourhood U of P and is an ϵ-approximation of it, with pre-assigned ϵ, if λ, μ are small enough (U being fixed). In particular, $F + \lambda G = H + \mu K = 0$ contains an arc C' approximating $C \cap U$.

9. First stage of approximation in 3-space. The lemmas of the preceding sections are now to be combined, the idea being to choose co-ordinates in such a way that the situations described in these sections all arise separately. Attention will be restricted meanwhile to closed piecewise algebraic curves.

Let C be a closed piecewise algebraic curve in 3-space with arcs C_i and joints P_j and assume that the tangents to the two arcs meeting at each joint are distinct. Assume that C_i is part of an algebraic curve $\bar{C}_i$ such that at each P_j just two branches belonging to these algebraic curves meet. Choose co-ordinates as follows.

(1) The projection of $\bar{C} = \bigcup \bar{C}_i$ on the (x, y)-plane is to be one-one with the exception that some double points, projections of two simple points of $\bar{C}$, may be introduced. The two tangents at each such double point are to be distinct. In particular, the joints P_j are to project regularly.

(2) The (x, z)-plane is not to be parallel to the tangents of the C_i at the joints.

It will be assumed, in addition, that all joints have been adjusted so that case (2) as described in §7 applies at each; this adjustment itself yields an arbitrarily good approximation of the curve.

For convenience the following points on the (x, y) and (y, z)-planes will be called special:

(*a*) The projections of all singularities of $\bar{C} = \bigcup \bar{C}_i$ except the P_j;

(*b*) Double points of the projection which are projections each of two simple points;

(*c*) Projections of points where the tangent to $\bar{C}$ is parallel to the (x, z)-plane;

Then the following theorem gives a first approximation to C.

THEOREM 4. *Let $F(x, y) = 0$ and $H(y, z) = 0$ be the equations of the projections of $\bar{C}$ on the (x, y) and (y, z)-planes respectively. Let $G(x, y)$ and $K(y, z)$ be polynomials not vanishing at the projections of the P_i. Then, if G and K are arranged to have suitable signs at the projections of the P_i and if they vanish to sufficiently high order at all special points and if λ, μ are small enough, $F + \lambda G = 0$, $H + \mu K = 0$ is a curve $\bar{C}'$ with a circuit C' which is an arbitrarily good approximation of C smoothed at the P_i and otherwise singularity preserving with analytic equivalence of arbitrarily high order at the singularities.*

Proof. The plan of the proof is similar to that of Theorem 1. The P_i and all points projecting on special points are surrounded by sufficiently small neighbourhoods for the appropriate lemma to apply. Thus, P_i has a neighbourhood $U(P_i)$ such that $\bar{C}'$ is a smooth approximation of C', and arbitrarily close if λ, μ are small enough. If P projects on a special point it has a neighbourhood $U(P)$ in which $\bar{C}'$ is analytically equivalent to $F = H = 0$. That is to say, part of $\bar{C}' \cap U(P)$ is analytically equivalent to C and this implies the existence of an operator f mapping $C \cap U(P)$ into $\bar{C}'$ such that the distance of Q from $f(Q)$ is arbitrarily small if λ, μ are small enough (see Lemma 6.2). Thus, in the union of the $U(P)$, where P is either a joint of projects or a special point, the operator f is constructed and is a one-valued continuous mapping. $f(Q)$ is in all cases arbitrarily close to Q if λ, μ are small enough. It is required to extend f to all of C. C outside the $U(P)$ is made up of non-singular arcs. In sufficiently small neighbourhoods of these arcs it is not hard to see that C' is a union of non-singular arcs. The extension of f is then made as in Theorem 1.

It will be noticed that the success of the proof of the above theorem depends upon the possibility of choosing polynomials G and K vanishing to a sufficiently high order at the singularities of C and with the correct signs at the P_i (see §15).

10. Second stage of approximation in 3-space. In this section C is a circuit of a real algebraic curve $\bar{C}$ in 3-space. Co-ordinates can be chosen so that $\bar{C}$ projects on a curve $\bar{C}_0$ in the (x, y)-plane, the correspondence between these curves being one-one except that a finite number of points of $\bar{C}_0$ are each projections of a pair of simple points of $\bar{C}$. Such points of $\bar{C}_0$ are to be double points where two simple branches meet with distinct tangents. If $\bar{C}_0$ has the equation $F(x, y) = 0$ then $\bar{C}$ can be represented by equations $F = 0$, $z = f/g$, where f/g is a rational function of x and y defined except possibly at singular points of $\bar{C}_0$.

THEOREM 5. *C being as above there exists a real algebraic curve $\bar{C}'$ of which an isolated circuit C' is an arbitrarily good approximation of C, singularity preserving with analytic equivalence of arbitrarily high order at each singularity.*

Proof. Apply Theorem 3 to $\bar{C}_0$, thus obtaining a curve $\bar{C}_0'$ of which one circuit C_0' is an arbitrarily good singularity preserving approximation of C_0 with analytic equivalence of order greater than a pre-assigned integer at each singularity, while $\bar{C}_0' - C_0'$ is contained in some pre-assigned set. Let the equation of $\bar{C}_0'$ be $F'(x, y) = 0$ and consider the curve $\bar{C}'$ with equations $F' = 0$, $z = f/g$. In particular let C' be a circuit of $\bar{C}'$ projecting on C_0'.

Let P be a singularity of C and apply Lemma 6.1, F and F' playing the part of the f_i, f_i' of that lemma, and both F and F' of the Lemma being replaced by $gz - f$. If the analytic equivalence of C_0 and C_0' is of sufficiently high order around the projection of P, then the lemma quoted implies that C and C' are analytically equivalent at P, to an arbitrarily high order.

A similar argument at points P_1 and P_2 of C projecting on the same point of C_0, shows that C' is analytically equivalent to C around P_1 and P_2.

Finally, the continuity of f/g at simple points of C_0 ensures that the correspondence between C and C' is one-one and is in fact the required approximation. It is clear that C' is an isolated circuit, and in fact $\bar{C}' - C'$ can be made to lie in a pre-assigned set. The proof is thus complete.

11. The final approximation theorems in 3-space.

THEOREM 6. *Let C be a closed piecewise algebraic curve with arcs C_i. C_i is to be part of an algebraic curve $\bar{C}_i$, and at each joint of C two simple branches of the $\bar{C}_i$ are to meet with distinct tangents. Then there exists an algebraic curve $\bar{C}'$ with an isolated circuit C' which is an arbitrarily good singularity preserving approximation of C, smoothed at the joints, and with analytic equivalence of arbitrarily high order at the singularities.*

Proof. The idea of the proof has already been sketched in §5. By Theorem 4, approximate C with a circuit C^* of an algebraic curve. Then apply Theorem 5 to C^*.

THEOREM 7. *Let C be a closed piecewise analytic curve in 3-space. Then there exists an arbitrarily good singularity preserving approximation of C by an isolated circuit of a real algebraic curve.*

Proof. In a sufficiently small neighbourhood of each singularity, C can be replaced by an analytically equivalent algebraic arc (2). The remainder of C can be approximated by straight line segments joined end to end. Thus, C has been approximated by a closed piecewise algebraic curve and it is not hard to see that the condition imposed in Theorem 6 can be assumed to be satisfied. The result follows at once from that theorem.

COROLLARY. *A similar result holds for an open piecewise analytic curve, for such a curve can always be closed by an auxiliary arc joining its end points.*

12. Approximation of a piecewise algebraic curve in n-space. An approximation theorem of this type has already been obtained for $n = 3$. The general result will be obtained by induction. Let C be a closed piecewise algebraic curve in n-space. C is part of a composite algebraic curve $\bar{C}$. Coordinates are to be chosen in such a way that $\bar{C}$ projects on the hyperplane $x_n = 0$ in a one-one manner and also in such a way that no tangent to $\bar{C}$ is parallel to the x_n-axis. Thus, under this projection no fresh singularities are introduced. If the arcs of C are denoted by C_i, C_i being part of a real algebraic curve $\bar{C}_i$, then it will be assumed that the joints P_j of C are the only points common to the $\bar{C}_i$. It will be remembered that a similar restriction was imposed on curves in 3-space but was eventually removed in the proofs of the approximation theorems.

Let K be the projection of C, $\bar{K}$ that of $\bar{C}$, K_i that of C_i and $\bar{K}_i$ that of $\bar{C}_i$ on $x_n = 0$. Then a point $(x_1, x_2, \ldots, x_n)$ belongs to $\bar{C}$ if and only if $(x_1, x_2, \ldots, x_{n-1})$ is on $\bar{K}$ and $x_n = f(x_1, x_2, \ldots, x_{n-1})$, where f if a continuous function which is rational on each $\bar{K}_i$ separately, projection being a birational mapping on each $\bar{C}_i$. The approximation of C is to be made by approximating K, by the induction hypothesis, in $x_n = 0$ and at the same time approximating f by a rational function F such that $|f - F|$ is small except at singularities of

$\bar{K}$. Near such a singularity the numerator and denominator of F are to differ by terms of arbitrarily high order from those of f.

Attention will now be fixed on approximation by rational functions of the type just indicated. Let A be a bounded closed set in Euclidean space. A real valued function f on A is called quasi-rational on A if there exists a finite set S of points of A, to be called the singularities of f, such that f is continuous on $A - S$ and such that there is a polynomial ψ vanishing at each point of S but at no other point of A and having the property that $f\psi$ is equal to a polynomial in some neighbourhood of each point of S.

The function F on A is called a rational approximation of the quasi-rational function f if;

(1) F is rational;

(2) Outside prescribed neighbourhoods of the points of $S, f - F$ is less than a pre-assigned number ϵ;

(3) If $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is in S, $f = \phi/\psi$ around $(\bar{x})$, ϕ being a polynomial and $F = \Phi/\Psi$, the fractions not necessarily being in lowest terms, then $\phi - \Phi, \psi - \Phi$ have orders greater than a pre-assigned integer r in the $x_i - \bar{x}_i$.

In the notation of **(3)** the approximation is said to be of order $> r$ at the singularity $(\bar{x})$. If, in addition, f is one-valued (as in the above situation of a function of a curve) then the inequality $|f - F| < \epsilon$ will be required to hold on all of A and the accuracy of the approximation can be specified by ϵ and r.

THEOREM 8. *Let f be a quasi-rational function on the closed bounded set A, and let $P_1, \dots, P_n$ be the set of singular points. Let U_i be a neighbourhood of P_i and let ϵ be a pre-assigned number. Then there exists a rational approximation of f approximating to within ϵ outside the U_i, and approximating to a pre-assigned order at the P_i.*

Proof. At P_i, f can be written as a rational function with denominator ψ; say $f = \phi_i/\psi$ in a neighbourhood of P_i. f is continuous outside the U_i, and so in particular ψf is continuous outside U_i.

Construct a polynomial Φ vanishing at the P_i in a similar way to the ϕ_i. A convenient method is as follows. Let g_i be a polynomial vanishing to order r at $P_j (j \neq i)$ and such that $g_i - 1$ vanishes to order r at P_i (cf. §15). Then set

$$\Phi(x) = \sum g_i \phi_i.$$

Now $\Phi - \psi f$ is continuous on $A - \bigcup U_i$, and so has a polynomial η_1-approximation G there. Let H be a rational function vanishing to order r at the P_i, and such that $1 - \eta_2 < H < 1$ outside the U_i (Lemma 3.5). Let $\phi = \Phi - GH$. Then

$$\begin{aligned} |\phi - \psi f| &= |\Phi - GH - \psi f| \\ &= |\Phi - \psi f - G + G(1 - H)| \\ &< \eta_1 + |G|\eta_2, \end{aligned}$$

outside the U_i. Hence, outside the U_i, $|\phi/\psi - f| < (\eta_1 + |G|\eta_2)/|\psi|$.

On $A - \bigcup U_i$ ψ is bounded below; and so, if η_1, η_2 are small enough, this last quantity is $< \epsilon$. Now compare the fractions ϕ/ψ and ϕ_i/ψ around P_i. They have the same denominators and the difference of their numerators is

$$\begin{aligned}\phi - \phi_i &= \sum g_j\phi_j - GH - \phi_i \\ &= \sum_{j \neq i} g_j\phi_j + (g_i - 1)\phi_i - GH\end{aligned}$$

and the terms of this expression are all of order not less than r at P, by the mode of definition of the g_i and of H.

To apply the above result to the approximation of the curve C some preliminary adjustments may be necessary. Suppose that $\bar{C}_i$ is given by

$$x_n = \frac{f_i(x_1, x_2, \ldots, x_{n-1})}{g_i(x_1, x_2, \ldots, x_{n-1})},$$

where $(x_1, x_2, \ldots, x_{n-1})$ is on $\bar{K}_i$. Suppose that g_1 vanishes at some of the joints of the K_i, say $P_j, P_k, \ldots$. These points will not include the ends of of K_1 since projection is one-one and regular at these end points. Let $h_1(x_1, x_2, \ldots, x_{n-1})$ be a polynomial vanishing on K_1 but not at $P_j, P_k, \ldots$ (by hypothesis these are not on $\bar{K}_1$). Then $f_1/(g_1 + ch_1) = f_1/g_1$ on K_1, c being a constant, and the denominator is not 0 at any P_i. If necessary, a similar adjustment is to be made for all the g_i.

Now set the fractions f_i/g_i with adjusted denominators over a common denominator g and rewrite as f_i/g. Then f defined as f_i/g on K_i is a quasi-rational function on K whose singular points, namely, the zeros of g, are all different from the P_i. The following approximation theorem can now be proved.

THEOREM 9. *Let C be a closed piecewise algebraic curve in n-space. Then there exists an arbitrarily good singularity preserving approximation of C by an isolated circuit of an algebraic curve, with smoothing at the joints and analytic equivalence of arbitrarily high order at the singularities.*

Proof. The result is true for $n = 3$ and will now be proved by induction. Assume that it is true for $n - 1$. Then, in the notation introduced at the beginning of the section, K has an approximation K' by an isolated circuit of an algebraic curve with the analytic equivalence at all singularities of the quasi-rational function f (all projections of singularities of C are to be included among these). Let F be a rational approximation of f.

Then, by Lemma 6.1, if the analytic equivalence of K and K' at singular points of f and at singularities of K is of sufficiently high order and if the approximating order of F to f at these points is sufficiently high also, then C is analytically equivalent to the curve $x_n = F(x_1, x_2, \ldots, x_{n-1})$ with $(x_1, x_2, \ldots, x_{n-1}) \in K'$, at the appropriate points of C. Apart from singular points it is clear that this curve is an approximation of C and it is certainly an isolated circuit of an algebraic curve.

The usual extension (similar to that made in §11) can be made here to closed and open piecewise analytic curves in n-space.

13. Approximation on a hypersurface. It is convenient to make a few remarks here on real algebraic varieties. A real algebraic variety V is the set of all real points on a complex algebraic variety V'. V' is understood to be contained in affine n-space over the complex numbers while V is a subset of Euclidean n-space. V' is to be chosen as the smallest complex algebraic variety containing V. Thus, V' has a simple point on V for otherwise V' could be replaced by its singular locus. Denote by $\bar{V}'$ the variety whose points are obtained from those of V' by taking the complex conjugates of all co-ordinates. Then it can be assumed that $\bar{V}' = V'$. For otherwise V' could be replaced by $\bar{V}' \cap V'$. The equations of V' can therefore be chosen to have real coefficients.

It is known that if V' is of dimension r then co-ordinates can be chosen in such a way that the equations of V' are of the form $f(x_1, x_2, \ldots, x_{r+1}) = 0$, $x_{r+1+i} = f_i(x_1, x_2, \ldots, x_{r+1})$ where f is a polynomial and the f_i $(i = 1, 2, \ldots, n - r - 1)$ are rational functions of their arguments. By the arguments made above it can also be arranged that the coefficients appearing in f and the f_i are all real numbers. It is then not hard to see that if P is a real simple point of V' then there are uniformising parameters whose real parts are real local co-ordinates, in the sense of real analytic manifolds, on V around P. That is to say, P has a neighbourhood analytically homeomorphic to a Euclidean r-cell. P is then called a simple point of V. Also the dimension of V is defined by the dimension of V', namely, r.

Now let C be a closed or open piecewise algebraic curve on a real algebraic hypersurface H in n-space. Assume that C is not contained in the singular locus of H and that the joints of C are all simple on H. Assume for the moment that $n > 3$. Choose co-ordinates so that the following conditions hold:

(1) H has a polynomial equation $F = 0$ and C is not contained in the locus with equations $F = \partial F/\partial x_n = 0$ and in particular the joints of C are not in this locus.

(2) C projects in a one-one manner on a curve K in the space $x_n = 0$.

THEOREM 10. *There exists an arbitrarily good singularity preserving approximation of C by a circuit or arc (the latter if C is open) of an algebraic curve on H, smoothed at the joints and with analytic equivalence of arbitrarily high order at the singularities.*

Proof. A smoothing approximation K' which is singularity preserving is to be constructed for K. Let $Q_1, Q_2, \ldots, Q_m$ be the singularities of K along with all the intersections of K with the projection of $F = \partial F/\partial x_n = 0$. Then K

and K' are to be made analytically equivalent at all the Q_i (Theorem 9). If this approximation is close enough and if the analytic equivalences just mentioned are of high enough order, then Lemma 6.1 implies analytic equivalence of C and part of the curve C' given by $F = 0$, with $(x_1, x_2, \ldots, x_{n-1})$ on K' around the Q_i. It is easy to see that these local approximations can be extended to the required approximation; the details of the argument are similar to those of the proof of Theorem 1. In particular, the smoothing of the approximation K' of K lifts into H since $F = 0$ can be solved for x_n around the points in question.

In the case $n = 3$ it cannot be assumed that the projection on $x_3 = 0$ is one-one on C. New singularities may be introduced. In the approximation of K by K', we can also make these curves analytically equivalent at any such new singularities. The lifting into H is carried out as before with the aid of Lemma 6.1. The theorem is thus proved for all values of n.

It is clear from the proof that C' and C can be made analytically equivalent at any further finite set of points in addition to those projecting on the Q_i.

Theorem 10 can be extended at once to the approximation of piecewise analytic curves on a hypersurface. For, let C be such a curve on the hypersurface H, satisfying a condition similar to that imposed in Theorem 10, namely, that no arc of C lies in the singular locus of H, and in particular, the joints of C are simple on H. Subdivide C into arcs such that on each of them the projection on the hyperplane $x_n = 0$ is one-one. Then apply the method of Theorem 10 to approximate each of these arcs by an algebraic arc on H, with analytic equivalence at all the singularities of C and also at all points of intersection of C with the locus having the equations $F = 0$, $\partial F/\partial x_n = 0$.

14. Approximation on a variety of any dimension.

THEOREM 11. *Let C be a piecewise analytic curve on a real algebraic variety V having at most a finite number of points in common with the singular locus of V. In particular the joints of C are to be simple on V. Then there exists an algebraic curve on V with a circuit or arc approximating C arbitrarily closely, smoothed at the joints and otherwise singularity preserving, with analytic equivalence of arbitrarily high order at the singularities.*

Proof. Suppose that V is of dimension n and is contained in $(n + r)$-space. The result is then known, by the last section, for $r = 1$, and the object is to prove it in general by induction on r.

Choose co-ordinates so that the equations of V are $f(x_1, x_2, \ldots, x_{n+1}) = 0$, where f is a polynomial, along with $x_{n+i} = f_i(x_1, x_2, \ldots, x_{n+1})$, $(i = 2, \ldots, r)$, where the f_i are rational functions of their arguments. That such a choice of co-ordinates can be made is a well-known theorem of algebraic geometry. There is thus a sequence of varieties $V_1, V_2, \ldots, V_r$, where V_s is contained in Euclidean $(n + s)$-space in which the co-ordinates are $x_1, x_2, \dots, x_{n+s}$, V_1 has the equation $f = 0$, and V_{s+1} projects on V_s in such a way that the points of V_{s+1} are of the form $(x_1, x_2, \dots, x_{n+s+1})$ with $(x_1, x_2, \dots, x_{n+s}) \in V_s$ and $x_{n+s+1} = f_{s+1}\ (x_1, x_2, \dots, x_{n+1})$. The curve C on $V = V_r$ projects on a curve C_s on V_s. Thus, the points of C_s are defined by $x_{n+s} = f_s(x_1, x_2, \dots, x_{n+1})$ with $(x_1, x_2, \dots, x_{n+s-1})$ on C_{s-1}. It can also be assumed that the co-ordinates are chosen so that the f_t are indeterminate at only finitely many points of C_s and that none of these is a joint of C_s.

By the induction hypothesis C_s can be approximated by a circuit or arc C_s' of an algebraic curve on V_s with smoothing at the joints, otherwise singularity preserving with analytic equivalence of arbitrarily high order at the singularities of C_s and also at all points of C_s at which any of the f_t is indeterminate. Define C_{s+1}' as the curve whose points are $(x_1, x_2, \dots, x_{n+s+1})$ with $(x_1, x_2, \dots, x_{n+s}) \in C_s'$ and $x_{n+s+1} = f_{s+1}(x_1, x_2, \dots, x_{n+1})$. Then C_{s+1}' is an arc or circuit of an algebraic curve on V_{s+1}. It is required to prove that it is an approximation of C_{s+1} with smoothing at the joints and otherwise singularity preserving. Clearly it is an approximation outside neighbourhoods of the following points: (a) singularities of C_{s+1}, (b) points of C_{s+1} singular on V_{s+1}, (c) points of C_{s+1} projecting on points of C_s at which some f_t is indeterminate. That C_{s+1}' is analytically equivalent to C_{s+1} around all such points follows at once from Lemma 6.1. The inductive proof is thus complete.

15. Some special polynomials. In this section explicit constructions are given for polynomials satisfying certain special conditions, such as were required in some of the proofs earlier in this paper.

The first such polynomial is to be a polynomial $F(P; Q; x)$ in the co-ordinates $x_1, x_2, \dots, x_n$ in n-space vanishing to the order r at P and such that $1 - F$ vanishes to the order r at Q. For convenience in defining this polynomial take P as the origin and let the co-ordinates of Q be $(x'_1, x'_2, \dots, x'_n)$. Then the definition is to be

$$F(P; Q; x) = 1 - \frac{\sum(x'^{r}_i - x^{r}_i)}{\sum x'^{r^2}_i}.$$

The next definition is to be that of a polynomial $F(P_1, P_2, \dots, P_n;\ Q;\ x)$ vanishing to the order r at $P_1, P_2, \dots, P_m$ and such that $1 - F$ vanishes to the order r at Q. A suitable definition for a polynomial with this property is the product of the $F(P_i; Q; x)$ for $i = 1, 2, \dots, m$, using the definition just given for the individual factors.

Finally, a polynomial is to be constructed which vanishes to the order r at the points $P_1, P_2, \dots, P_m$ and has given values $k_1, k_2, \dots, k_p$ at another set of points $Q_1, Q_2, \dots, Q_p$. A suitable definition for such a polynomial is

$$\sum k_i F(P_1, P_2, \dots, P_m, Q_1, Q_2, \dots, \hat{Q}_i, \dots, Q_p; Q_i; x)$$

where the circumflex denotes the omission of the letter marked.

PART III: SHEETS OF A REAL ALGEBRAIC VARIETY

16. Definition and examples. A subset S of a real algebraic variety V will be called analytically connected if every pair of points of S can be joined by an analytic arc contained entirely in S. A subset S of V is called a sheet of V if S is analytically connected and is not contained in any larger analytically connected set on V.

This definition is slightly weaker than that given by Nash. The term "sheet" in **(1)** is equivalent to the term "proper sheet" according to the following definition.

The sheet S of V is called proper if there is a point of S with a neighbourhood U such that $U \cap V \subset S$. If this condition is not satisfied S will be called embedded.

Examples. (1) Consider the surface in 3-space with the equation $(y^2 + z^2)^2 = z^2x^3$. The cross-section of this surface parallel to the (y, z)-plane for $x > 0$ consists of two circles touching while for $x < 0$ the only real points are in the x-axis. The two circles referred to have equal radii proportional to $x^{3/2}$. It is not hard to see that this surface has two sheets. One is the x-axis and the other is the part of the surface with $x \geqslant 0$. The first statement is clear since the x-axis is analytically connected and any analytic arc on the surface through a point with $x < 0$ must lie entirely on the x-axis. To prove the second statement take any points P and Q on S with $x \geqslant 0$ and project them on the (y, z)-plane. Let the projections be P' and Q'. If these points are on the same side of $z = 0$, join them by a straight line with parametric equations $z = z(t)$, $y = y(t)$. Then

$$x(t) = \frac{(y^2(t) + z^2(t))^{2/3}}{z(t)^{2/3}}$$

is real analytic and gives the required arc on the surface joining P and Q. If P' and Q' are on opposite sides of $z = 0$ join them by an analytic arc $y = y(t)$, $z = z(t)$ such that the origin corresponds to $t = 0$ and such that this arc crosses $z = 0$ only at the origin. Assume that around $t = 0$ y and z have power series expansions $y(t) = at + \ldots$, $z(t) = bt + \ldots$ with a and b non-zero. Then $x(t)$ is an analytic function of s where $t = s^3$. This again gives an analytic arc on the surface joining P and Q, all points of the arc satisfying $x \geqslant 0$. It should be noted that when P and Q are on opposite sides of $z = 0$ then an analytic arc joining them must necessarily pass through the origin. In this example both the sheets occurring are proper.

(2) Consider now the surface S

$$(y^2 + z^2)^4 - z^4x^6 + (y^2 + z^2)^r = 0$$

where $r \geqslant 8$. Let $F = (y^2 + z^2)^4 - z^4x^6$. Then

$$-\frac{2}{3}x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z}$$

is a constant multiple of $(y^2 + z^2)^4$ and so for $r > 8$, $(y^2 + z^2)^r$ is in the square of the ideal generated by F_x, F_y, and F_z. The theorem of Samuel **(1)** shows that the surface S near the origin is shaped like two copies of the surface in example (1) placed point to point. On the other hand, the equation of S can be written as

$$x^2z^2 = \pm(y^2 + z^2)^2(1 + (y^2 + z^2)^{r-4})^{\frac{1}{2}}.$$

The procedure of the example of (1) shows that the sets $x > 0$ and $x < 0$ on this surface are analytically connected. Also, the second method of writing the equation of S shows that no analytic arc connects points with $x > 0$ to points with $x < 0$. The two sets on S given by $x > 0$ and by $x < 0$ are thus two separate sheets and are proper. The x axis is also a sheet for it is analytically connected and not contained in either of the above sheets. It is embedded.

17. Local dimension of a sheet or variety. Let V as before be a real algebraic variety contained in the complex algebraic variety $V' = \bar{V}'$ of complex dimension n. It will be convenient to speak of n as being the dimension of V. Let S be a sheet of V and let p be a point of S.

The local dimension of S at p, written as $\dim_p S$, will be said to be equal to n if every neighbourhood of p contains a simple point of V lying on S. Otherwise $\dim_p$ will be said to be less than n.

THEOREM 12. *Let S be a sheet of the real n-dimensional algebraic variety V and let p be a point of S. If* $\dim_p S = n$, *then* $\dim_q S = n$ *for all q on S.*

Proof. Since $\dim_p S = n$, a neighbourhood U of p will contain a simple point p' of V lying on S. If q is on S then there is an analytic arc A on S joining p' and q. This arc is not entirely contained in the singular locus of V since p' is simple, and so it meets this locus at a finite number of points. The last statement is equivalent to the fact that an analytic function of t for $0 < t < 1$ has only a finite number of zeros. Therefore, there is a simple point of V on S, namely, on A, in any neighbourhood of q. Therefore, $\dim_q S = n$.

COROLLARY. *If* $\dim_p S < n$ *for some p on S then S consists entirely of singular points of V.*

Proof. For if q is a simple point of V lying on S then $\dim_q S = n$ and so, by the above theorem, $\dim_p S < n$ is impossible.

If $\dim_p S = n$ for some p on S the above theorem justifies defining the dimension of S by n. In the contrary case the dimension of S will be said to be less than n.

Now, in the case where the dimension of S is less than n, S is contained in the singular locus of V. Thus, S is a sheet of a subvariety of V. Let V_0 be the smallest subvariety of V containing S and let the dimension of V_0 be r. The above theorem and its corollary show that $\dim_p S$, S being regarded as a sheet of V_0, is r at each point p of S. That is to say, the dimension on S is r. For otherwise S would be contained in the singular locus of V_0, namely, a smaller subvariety than V_0.

Thus, the dimension of S can be defined in all cases as the dimension of the smallest real algebraic variety containing S.

Some properties of n-dimensional sheets of n-dimensional varieties can be deduced from the following semi-transitivity property of analytic connectivity.

LEMMA 17.1. *Let p, q, r be points on V, q being simple. Then, if there are analytic arcs on V joining p to q and q to r, there is also an analytic arc on V joining p to r and meeting a pre-assigned neighbourhood of q.*

Proof. The union of the two arcs joining p to q and q to r is a piecewise analytic arc on V. By Theorem 11 there is an algebraic arc approximating it arbitrarily closely, smoothing at q and otherwise singularity preserving. This give the required joint of p and r. Note incidentally that p and r may be singular on the given arcs, and these singularities must be preserved along with any others.

THEOREM 13. *Let V be a real algebraic variety of dimension n and let p be a simple point. Let S be the set of all points joined by analytic arcs to p on V. Then S is an n-dimensional sheet of V and every sheet of dimension n can be obtained in this way.*

Proof. Let q_1 and q_2 be points of S. Then there are analytic arcs on V joining q_1 to p and p to q_2. Let U be a neighbourhood of p homeomorphic to an n-cell. Lemma 17.1 implies that there is an analytic arc on V joining q_1 to q_2 and meeting U in some point q, say q being a simple point of V. Take q' on the arc q_1q_2. Then there is an analytic arc joining q' to q on V, namely, part of the arc q_1q_2, and there is also an analytic arc in the cell U joining q to p. Applying again Lemma 17.1 it follows that there is an analytic arc on V joining q' to p. Therefore, q' belongs to S and so the whole arc q_1q_2 lies in S. That is to say, it has been shown that S is analytically connected.

It must be shown now that S is a maximal analytically connected set. Assume that $S \subset S'$ where S' is analytically connected. If q is a point of S' there exists an analytic arc joining p and q in S' and so in V. It follows that $S' \subset S$ and the maximality of S is established.

Obviously $\dim_p S = n$ and so S is n-dimensional.

Conversely, let S be an n-dimensional sheet of V. Then S contains by definition a simple point p of V. Every point of S can be joined to p by an analytic arc lying in S and so lying in V. The above result and the maximal property of S show that S is the set of all points which can be joined to p in V by analytic arcs.

COROLLARY 1. *Each simple point of V belongs to exactly one sheet.*

COROLLARY 2. *Each n-dimensional sheet of a real n-dimensional variety V is proper.*

Proof. If S is of dimension n there is a simple point p of V on S. It follows that there is a neighbourhood U of p such that $U \cap V$ is an n-cell. All points of $U \cap V$ can be joined to p by analytic arcs on V and so $U \cap V$ lies in the sheet determined as in the above theorem by p. This sheet must be S and so S is proper.

The notion of local dimension can also be introduced for a real algebraic variety V (and, in fact, more generally for any real algebroid variety). If p is a point of V then the local dimension of V at p, written $\dim_p V$, will be said to be n if every neighbourhood of p contains a simple point of V, that is to say, a real simple point of V' in the terminology of §13. Otherwise $\dim_p V$ will be said to be less than n.

If $\dim_p V < n$ then there is a subvariety V_0 of V consisting entirely of singular points and there is a neighbourhood U of p such that $U \cap V_0 = U \cap V$. Let V_0 be the smallest real subvariety of V with this property. Then every neighbourhood of p must contain a simple point of V_0; for otherwise V_0 could be replaced by its singular locus, a smaller subvariety. If V_0 is of dimension r then $\dim_p V_0 = r$. Define now $\dim_p V = \dim_p V_0$.

Note that a variety is not homogeneous with respect to the notion of local dimension, whereas a sheet of a variety is. For example, on the surface of example (1) in §16 points satisfying $x > 0$ have local dimension 2 whereas those satisfying $x < 0$ have local dimension 1.

18. Local study of a real algebraic variety. To get further information of the sheets of a variety some results on the local structure of a real algebraic variety are required. These will be obtained in the following three lemmas.

LEMMA 18.1. *Let p be a point of a real algebraic variety V in n-space. Then, in any pre-assigned neighbourhood of p there is a neighbourhood U which can be written as the union of the closures of a finite number of disjoint open n-cells U_i such that the union of the frontiers of the U_i is of the form $W \cap U$ where W is a real algebraic variety containing V. In addition, each U_i has p on its frontier.*

Proof. The proof will be carried out by induction on n. Assume first that $\dim_p V = n - 1$. Take p as origin and choose co-ordinates in such a way that V, which is a hypersurface, has an equation of the form

$$F = x_n^r + a_1 x_n^{r-1} + \ldots + a_r = 0,$$

where the a_i are analytic in $x_1, x_2, \ldots, x_{n-1}$ at p. This simply means that the x_n-axis does not lie in V. Let V_0 be the projection on $x_n = 0$ of the locus with equations $F = \partial F/\partial x_n = 0$. The induction hypothesis implies that there is a neighbourhood U_0 of p in $x_n = 0$ such that U_0 can be written as $\bigcup \bar{Z}_i$, where the Z_i are disjoint open $(n-1)$-cells and $\bigcup \mathrm{Fr} Z_i = U_0 \cap W_0$, where W_0 is a variety containing V_0. For each Z_i there are two possible cases to consider.

(1) There are sets on V, say $Z_i^{(1)}, Z_i^{(2)}, \ldots, Z_i^{(s)}$, projecting homeomorphically on Z_i and having p in their closures.

(2) There are no such sets as in (1).

Let C be a cylindrical neighbourhood of p, specified as the set of all points $(x_1, x_2, \ldots, x_n)$ with $(x_1, x_2, \ldots, x_{n-1})$ in some neighbourhood of p and x_n satisfying an inequality of the type $|x_n| < k$. C can be chosen as follows. If Z_i' is a set on V projecting homeomorphically on Z_i presenting case (2) or if Z_i' projects on a set Z_i presenting case (1) but is different from $Z_i^{(1)}, Z_i^{(2)}, \ldots, Z_i^{(s)}$, then $C \cap Z_i' = \phi$. Also, C is to be taken so that the subsets $x_n = \pm k$ of C do not meet V. This choice is always possible since the x_n-axis does not lie in V.

Shrink U_0 if necessary so that $U_0 \subset C$; this can be done by the induction hypothesis. Then define U as the set of points $(x_1, x_2, \ldots, x_n)$ such that $(x_1, x_2, \ldots, x_{n-1}) \in U_0$, $|x_n| < k$ for some positive number k. The cell decomposition of U is now to be defined. The part of U over a set Z_i presenting case (1) is divided into open cells by the $Z_i^{(j)}$. On the other hand, the part of U over a set Z_i presenting case (2) is itself an n-cell. Define the U_i as the collection of all these cells. It is at once clear that the U_i are disjoint and that p is in $\bar{U}_j$ for each j.

The union of the frontiers of the U_j consists of $V \cap U$ along with the top and bottom of U and the subset of U projecting on $\bigcup \mathrm{Fr} Z_i$. The last set can be written as $U_0 \cap W_0$ where W_0 is a real algebraic variety, by the induction hypothesis. Therefore, $\bigcup \mathrm{Fr}\, U_j$ is of the form required by this lemma. Also U can be taken arbitrarily small and so the proof is complete if $\dim_p V = n - 1$.

If $\dim_p V < n - 1$, repeat the above proof with V_0 taken as the projection of V on $x_n = 0$. Here only the sets Z presenting case (2) will appear but the rest of the proof is as above.

Lemma 18.2. *Let p be a point on a real algebraic variety V of dimension n and let W be a subvariety of V containing p. In any pre-assigned neighbourhood of p there is a neighbourhood U of p such that $V \cap U$ is the union of the closures of a set of disjoint open cells of dimensions $\leqslant n$ such that:*

(1) *$\bigcup \mathrm{Fr}\, U_i = U \cap W'$ where W' is a variety on V containing W.*

(2) *Each r-cell in the decomposition of $V \cap U$ is contained in exactly one proper sheet of V of dimension r.*

(3) *$p \in \bar{U}_i$ for each i.*

Proof. Note first that Lemma 18.1 is the special case of this lemma with V replaced by n-space. The general proof will be carried out by induction, the result being obvious for a curve. Assume that the theorem is true for any variety V such that $\dim_p V < n$ in any space. The result is then to be proved for a variety of local dimension n at p. The proof will first be carried out for a variety V in $(n + 1)$-space with $\dim_p V = n$. V must thus be a hypersurface and so co-ordinates can be chosen so that it has an equation of the form

$$F = x_{n+1}^r + a_1 x_{n+1}^{r-1} + \ldots + a_r = 0$$

where the a_i are analytic at p which is to be taken as origin. Project on $x_{n+1} = 0$ and let W_0 be the union of the projections of W and of the locus with equations $F = \partial F/\partial x_{n+1} = 0$. Apply Lemma 18.1 and use the notation used there. Then there is an arbitrarily small neighbourhood U_0 of p in $x_{n+1} = 0$ such that $U_0 = \bigcup \bar{Z}_i$, where the Z_i are disjoint open n-cells the union of whose frontiers is a variety W_1 containing W_0. In the terminology of Lemma 1, if Z_i presents case (1) there exists a finite number of sets $Z_i^{(j)}$ on V projecting homeomorphically on Z_i, p lying in the closure of each of them. Let U be chosen as in Lemma 18.1 and let $q \in V \cap U$. Then there are two cases to consider according as $\dim_q V = n$ or $\dim_q V < n$.

If $\dim_q V = n$, every neighbourhood N of q contains a simple point q' of V. Then, in a suitable neighbourhood N' of q' contained in N, there is a point q'' which is simple on V and does not project on the variety W_1 which contains the frontiers of Z_i. Then a neighbourhood of q'' projects homeomorphically into a subset of some Z_i. That is to say, q'' is in some set Z_i' projecting homeomorphically on Z_i. It follows at once, since N is any neighbourhood of q, that q is in the closure of Z_i'. By the choice of U, namely, as in Lemma 18.1, Z_i' must be one of the $Z_i^{(j)}$ having p in its closure. Hence all points q of $V \cap U$ with $\dim_q V = n$ are in the closure of some $Z_i^{(j)}$.

All points q in $V \cap U$ with $\dim_q V < n$ are contained in a subvariety V_0 of V. Apply the induction hypothesis to V_0, shrinking U if necessary. Thus, $V_0 \cap U$ is the union of the closures of a number of cells which, if taken along with the $Z_i^{(j)}$ provide the required cell decomposition of V.

The conditions (1) (2) (3) of the theorem must now be checked. Condition (1) follows from the induction hypothesis on V_0 and from the mode of construction of $Z_i^{(j)}$; (3) follows in the same way. Now (2) will be checked. $Z_i^{(j)}$ lies on exactly one proper n-dimensional sheet of V, namely, that determined by any simple point on it (Theorem 13). Let U_1 be one of the open cells of the decomposition of V_0 assumed in the induction hypothesis. Then, by this hypothesis, U_1 is contained in exactly one proper sheet S of V_0. If S is a proper sheet of V the result is proved. Suppose S is not proper. Then every neighbourhood of every point of S contains points of V not in S. Such points are also not in V_0, since S is proper in V_0. V therefore has local dimension n at such points and so the cell U_1 can be discarded, being contained in one of the $\bar{Z}_i^{(j)}$.

The proof is thus complete for a variety V with $\dim_p V = n$ contained in $(n + 1)$-space. To prove the result for a variety V of dimension n in $(n + r)$-space project V into $(n + 1)$-space. Let V_1 be the projection and let W_1 be the union of the projection of W and the variety of all points which are the projections of more than one point of V. Apply the result already obtained to V_1 with the subvariety W_1 and lift the cell decomposition so constructed to V.

Lemma 18.3. *Let V be a real algebraic variety, W a subvariety, and p a point of W. Then there is a neighbourhood U of p such that all points of $U \cap (V - W)$ can be joined to p by analytic arcs on V meeting W only at p.*

Proof. The proof is to be carried out by induction on $\dim_p V$. Assume that the result is true for any variety whose local dimension is less than n; the theorem is obvious in the case of a curve. The proof will first be carried out taking V as n-space and W as any variety through p. There are two cases to consider.

Case (1), $\dim_p W < n - 1$. Project W on the hyperplane $x_n = 0$, the projection being W'. Let p' be the projection of p. Apply the induction hypothesis taking V as the $(n-1)$-space $x_n = 0$ and replacing W by W'. Then there is a neighbourhood U' of p' such that all points of $U' - W'$ can be joined to p' by analytic arcs meeting W' only at p'. Also apply the induction hypothesis with V, W replaced respectively by W_1, W where W_1 is the set of all points projecting on W'. Then there is a neighbourhood U of p such that all points of $U \cap (W_1 - W)$ can be joined to p by analytic arcs in W_1 meeting W only at p. It can be assumed that U is so small that it projects inside U' and it can also be assumed to be cylindrical.

Let q be any point of $U - W$. If $q \in W_1$ there is an analytic arc in $U \cap (W_1 - W)$ joining p to q, meeting W only at p. On the other hand, if $q \notin W_1$, q projects on $q' \in U' - W'$ and so there is an analytic arc in U' joining p' and q' and meeting W' only at p'. This arc can clearly be lifted into an arc joining p and q and meeting W only at p. This completes the proof of the lemma with $V = n$-space in case (1).

Case (2), $\dim_p W = n - 1$. This time W is a hypersurface. Choose coordinates so that p is the origin and W has an equation of the form

$$F = x_n^r + a_1 x_n^{r-1} + \dots + a_r = 0$$

where the a_i are analytic in $x_1, x_2, \dots, x_{n-1}$ at p. Let W' be the projection on $x_n = 0$ of the locus with equations $F = \partial F/\partial x_n = 0$ and let W_1 be the set of points projecting on W'. Let $W_2 = W_1 \cap W$.

Apply the induction hypothesis with V, W replaced by W_1, W_2 respectively, thus obtaining a neighbourhood U of p such that all points of $U \cap (W_1 - W_2)$ can be joined to p by analytic arcs in W_1 meeting W only at p. Assume that U is cylindrical and is shrunk, if necessary, so that it has the properties of the neighbourhood U in Lemma 18.1. Let Z_i and $Z_i^{(j)}$ be as in that lemma. Apply the induction hypothesis with V, W replaced by the hyperplane $x_n = 0$ and W' respectively. Then there is a neighbourhood U' of p in $x_n = 0$ whose points can be joined to p by analytic arcs meeting W' only at p. Assume that U is shrunk, if necessary, so that it projects into U'. Let q be a point of $U - W$. If $q \in W_1$, it has been shown that there is an analytic arc joining p to q in $U \cap (W_1 - W)_2$ meeting W only at p. On the other hand, if $q \notin W_1$ and if q does not project on a set Z_i covered by the $Z_i^{(j)}$ then proceed as in case (1). If $q \notin W_1$ and q projects on Z_i covered by some of the $Z_i^{(j)}$ then there is an analytic arc in the interior of Z_i joining the projection q' of q to p and meeting the frontier of Z_i only at p. For the sake of definiteness assume that q lies between $Z_i^{(1)}$ and $Z_i^{(2)}$ and suppose that the above-mentioned arc from q' to p in Z_i has parametric equations

$$x_j = f_j(t), \qquad j = 1, 2, \ldots, n - 1.$$

Suppose that the points of $Z_i^{(1)}$ and $Z_i^{(2)}$ lying over this arc are given respectively by $x_n = f_n^{(1)}(t)$ and $x_n = f_n^{(2)}(t)$. Then the arc with equations

$$\begin{aligned} x_j &= f_j(t), \qquad j = 1, 2, \ldots, n - 1, \\ x_n &= hf_n^{(1)} + kf_n^{(2)}, \end{aligned}$$

for suitable h, k, is an analytic arc joining q to p in U meeting W only at p. This completes case (2).

The proof will now be carried out for any real algebraic variety V with $\dim_p V = n$. By Lemma 18.2 there exists a neighbourhood U_1 of p such that $U_1 \cap V$ is the union of the closures of disjoint open cells whose frontiers lie on a variety W_1 containing W. Also the proof of Lemma 18.2 shows that the n-cells in this decomposition project homeomorphically on n-cells in n-space, the frontiers of the latter being contained in the projection of W_1. All the cells in this decomposition whose dimensions are less than n lie on a variety W_2 and it will be assumed that W_2 contains W_1.

Apply the induction hypothesis to W_2 with the subvariety W. Then there is a neighbourhood U_2 on p such that all points of $U_2 \cap (W_2 - W)$ can be joined to p by analytic arcs in W_2 meeting W only at p. If $q \notin W_2$ then q is in the interior of an n-cell Z' projecting on an n-cell Z in n-space. Apply the result already proved for n-space with the subvariety W_3 which is the projection of W_1. Then there is a neighbourhood U_3 of the projection of p in n-space such that p can be joined by an analytic arc to any point of $U_3 - W_3$, and in particular to the projection of q. Such an arc meets the frontier of Z only at the projection of p and so can be lifted into Z'. With a suitable choice of parameter a lifted arc is still analytic at p. The neighbourhood U required by the statement of this lemma can be taken as the smallest of U_1, U_2, U_3.

19. Further properties of sheets.

THEOREM 14. *Each sheet S of a real algebraic variety V is a closed set.*

Proof. If V is of dimension n it is sufficient to prove the theorem for an n-dimensional sheet. For every other sheet is of maximal dimension in some subvariety which is itself a closed set of V.

If $p \in \bar{S}$ then every neighbourhood U of p meets S; let $q \in U \cap S$ and assume that U is open. U is a neighbourhood of q and S is n-dimensional and so, by definition, U contains a simple point q' of V, q' lying on S. By Lemma 18.3, there exists an analytic arc joining p to q' on V if U is small enough. It follows that p lies on the sheet of V determined by the simple point q' as in Theorem 13. By the corollary of that theorem this sheet is S. Since p is any point of $\bar{S}$ this shows that S is closed as required.

Corollary. *Every point p of a real algebraic variety belongs to some proper sheet.*

Proof. By Lemma 18.2 (2), there is a neighbourhood U of p which can be written as the union of the closures of open cells each of which is contained in some proper sheet. The point p is in the closure of each such cell and so is in the closure of some proper sheet. By the theorem just proved p lies on that sheet.

Theorem 15. *The number of sheets of a real algebraic variety in Euclidean space is finite.*

Proof. It is sufficient to prove the theorem for sheets of dimension n of an n-dimensional variety V because all other sheets are contained in some subvariety. Assume that the variety V has infinitely many n-dimensional sheets. Take a point on each sheet. This set of points will have a limit point p which may be a point at infinity. In the latter case apply some transformation, for example, inversion in some hypersphere, to make the limit point finite. Then every neighbourhood of p meets infinitely many n-dimensional sheets of V. But, by Lemma 18.2 there exists a neighbourhood U of p such that $U \cap V$ is a finite union of closures of cells, each n-cell lying on exactly one n-dimensional sheet. Since the sheets are closed (Theorem 14), the closure of each of these n-cells lies on exactly one n-dimensional sheet. Therefore, U meets only a finite number of these sheets. The contradiction so obtained proves the theorem.

References

1. J. Nash, *Real algebraic manifolds*, Ann. of Math., *56* (1952), 405–421.
2. P. Samuel, *Sur l'algébricité de certains points singuliers*, J. de Math. pures et appl. (9), *35* (1956), 1–6.
3. A. H. Wallace, *Algebraic approximation of manifolds*, Proc. London Math. Soc. (3), *7* (1957), 196–210.
4. H. Whitney, *Elementary structure of real algebraic varieties*, Ann. of Math., *66* (1957), 545–556.

University of Toronto

ON CERTAIN PAIRS OF MATRICES WHICH GENERATE FREE GROUPS

BOMSHIK CHANG, S. A. JENNINGS AND RIMHAK REE

1. Introduction. Denote by $F_{\alpha,\beta}$ the multiplicative group generated by the two matrices

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix},$$

where α and β are complex numbers. Sanov **(3)** proved that $F_{2,2}$ is free, and Brenner **(1)** showed that $F_{m,m}$ is free if $m \geqslant 2$.

In this note we extend these results, proving that $F_{\alpha,\beta}$ is free if $\alpha\beta$ satisfies all three of the conditions

$$|\alpha\beta| \geqslant 2, |\alpha\beta - 2| \geqslant 2 \quad \text{and} \quad |\alpha\beta + 2| \geqslant 2$$

(Theorem 2). In § 3, we prove that the set of algebraic numbers $\alpha\beta$ for which $F_{\alpha,\beta}$ is free is dense in the whole complex plane, and exhibit some values of $\alpha\beta$ for which $F_{\alpha,\beta}$ is not free. In the last section we show that the main idea used in the proof of Theorem 2 can be applied to a more general case.

2. Main theorems.

THEOREM 1. *If $\alpha\beta = \gamma\delta \neq 0$ then $F_{\alpha,\beta}$ and $F_{\gamma,\delta}$ are isomorphic.*

Proof. It suffices to prove that $F_{\alpha,\beta} \cong F_{\alpha\beta,1}$. We have

$$\begin{pmatrix} 1 & \alpha\beta \\ 0 & 1 \end{pmatrix} = P^{-1}\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}P, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = P^{-1}\begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}P, \quad \text{where } P = \begin{pmatrix} 1 & 0 \\ 0 & \beta \end{pmatrix}.$$

Hence the mapping $X \rightarrow P^{-1}XP$ gives the required isomorphism.

By Theorem 1, the consideration of the general group $F_{\alpha,\beta}$ is reduced to that of $F_\lambda = F_{2,\lambda}$ where $\alpha\beta = 2\lambda$. We shall say that a complex number λ is *free* if the group F_λ is free.

THEOREM 2. *Any complex number λ which satisfies*

$$|\lambda| \geqslant 1, |\lambda - 1| \geqslant 1, |\lambda + 1| \geqslant 1 \tag{2.1}$$

is free.

In order to prove Theorem 2, we adopt the following notation: for a complex variable z and a matrix

Received December 30, 1956. This research was in part supported by National Research Council of Canada grant G. 489.

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ad - bc = 1,$$

with complex entries, we denote by $P(z)$ the number $(az + b)/(cz + d)$. Then, as is well known, $(QP)\ (z) = Q(P(z))$ where Q is another such matrix.

In the rest of this section, we set

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix},$$

and define point sets $\mathscr{D}$ and $\mathscr{D}'$ in the complex plane as follows:

$$\mathscr{D} = \{z \mid |\Re z| < 1\}, \quad \mathscr{D}' = \{z \mid |\Re z| > 1\},$$

where $\Re z$ denotes the real part of z.

The following lemma is well known:

LEMMA 1. *If $z \in \mathscr{D}'$ then*

$$|z^{-1} - \tfrac{1}{2}| < \tfrac{1}{2} \quad \textit{or} \quad |z^{-1} + \tfrac{1}{2}| < \tfrac{1}{2}.$$

If, on the other hand,

$$|z - \tfrac{1}{2}\,| > \tfrac{1}{2} \quad \textit{and} \quad |z + \tfrac{1}{2}| > \tfrac{1}{2},$$

then $z^{-1} \in \mathscr{D}$.

LEMMA 2. *If λ satisfies* (2.1), *then $z \in \mathscr{D}'$ implies $B^n(z) \in \mathscr{D}$ for any non-zero integer n.*

Proof. Since

$$B^n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & n\lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

we may write

$$z_1 = z^{-1}, z_2 = z_1 + \mu, B^n(z) = z_2^{-1},$$

where $\mu = n\lambda$, which clearly satisfies (2.1). By Lemma 1, $z \in \mathscr{D}'$ implies

$$|z_1 - \tfrac{1}{2}| < \tfrac{1}{2} \quad \text{or} \quad |z_1 + \tfrac{1}{2}| < \tfrac{1}{2}.$$

If $|z_1 - \frac{1}{2}| < \frac{1}{2}$, then

$$|z_2 - \tfrac{1}{2}| = |z_1 + \mu - \tfrac{1}{2}| > |\mu| - |z_1 - \tfrac{1}{2}| > 1 - \tfrac{1}{2} = \tfrac{1}{2},$$

and

$$|z_2 + \tfrac{1}{2}| = |z_1 + \mu + \tfrac{1}{2}| > |\mu + 1| - |z_1 - \tfrac{1}{2}| > 1 - \tfrac{1}{2} = \tfrac{1}{2}.$$

Similarly, if $|z_1 + \frac{1}{2}| < \frac{1}{2}$, then we have

$$|z_2 - \tfrac{1}{2}| > \tfrac{1}{2} \quad \text{and} \quad |z_2 + \tfrac{1}{2}| > \tfrac{1}{2}.$$

Then again by Lemma 1, we have $B^n(z) \in \mathscr{D}$ as required.

Proof of Theorem 2. We shall derive a contradiction by assuming that F_λ is not free. If F_λ is not free there must exist a non-trivial word G of F_λ such that

$$G = B^{n_r}A^{m_r} \dots B^{n_1}A^{m_1} = E \tag{2.2}$$

where we can always assume that the integers $m_1, n_1, \dots, m_r$ are all not zero. Define

$$\begin{aligned} &z_1 = A^{m_1}(0), && z_1' = B^{n_1}(z_1), \\ &z_k = A^{m_k}(z_{k-1}'), && z_k' = B^{n_k}(z_k), \qquad (k < r-1), \end{aligned} \tag{2.3}$$

and $\quad z_r = A^{-m_r}B^{-n_r}(0).$

Then (2.2) implies

$$z_{r-1}' = z_r. \tag{2.4}$$

Since $z_1 = 2m_1$, $|\Re z_1| > 1$, $z_1 \in \mathscr{D}'$. By Lemma 2, we have

$$z_1' = B^{n_1}(z_1) \in \mathscr{D}.$$

Then, by (2.3), $z_2 = z_1 + 2m_m$, and hence

$$|Rz_2| > |2m_2| - |\Re z_1'| > 1.$$

Thus, $z_2 \in \mathscr{D}'$. Again, using Lemma 2, we have

$$z_2' = B^{n_2}z_2 \in \mathscr{D}.$$

Repeating this argument, we find $z_{r-1}' \in \mathscr{D}$. On the other hand, $z_r = -2m_r$, and $|\Re z_r| > 2$. Therefore $z_r \notin \mathscr{D}$ and $z_{r-1}' \neq z_r$, which contradicts (2.4). Thus Theorem 2 is proved.

3. Distribution of free and non-free points. First we note that *any transcendental number is free.* This is seen as follows. Let

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix},$$

where x is an indeterminate. Then any word G generated by A and B is of the form

$$G = \begin{pmatrix} p_1(x) & p_2(x) \\ p_3(x) & p_4(x) \end{pmatrix},$$

where $p_1(x), \dots, p_4(x)$ are polynomials in x with integral coefficients. In the course of proving Theorem 2 we have shown that if G is of the form given in (2.2) and if λ is a number satisfying (2.1), then $G(0) = p_2(\lambda)/\mathrm{p}_4(\lambda) \neq 0$. Hence, for such an element G, the polynomial $p_2(x)$ is always not identically zero. Therefore $G \neq E$ for any transcendental value of x. From this it follows that any transcendental number is free.

The following theorem, however, shows that there are also many free algebraic numbers in the domain excluded by Theorem 2.

THEOREM 3. *Any point in the complex plane is a limit of algebraic free points.*

In order to prove Theorem 3, we need the following well-known result (4, p. 122).

LEMMA 3. *Let p be a prime and c a rational number. Then the polynomial $x^p - c$ is reducible over the rational field, if and only if, c is a* p*th power of a rational number.*

Proof of Theorem 3. Let w be a given complex number. Because of Theorem 2, we may assume that w lies in the domain excluded by Theorem 2. For any positive number ϵ there exist a prime p, an integer q, and a rational number a such that $|w - \lambda_1| > \epsilon$ and $\lambda_2 > 4$, where

$$\lambda_1 = a + 2^{2+1/p}e^{2q\pi i/p}, \lambda_2 = a + 2^{2+1/p}.$$

We shall show that λ_1 is free. Assume the contrary. Then there must be a non-trivial word G of $F_{2,x}$

$$G = \begin{pmatrix} p_1(x) & p_2(x) \\ p_3(x) & p_4(x) \end{pmatrix}$$

which becomes the identity matrix when $x = \lambda_1$. Hence, we have

$$p_1(\lambda_1) - 1 = p_2(\lambda_1) = p_3(\lambda_1) = p_4(\lambda_1) - 1 = 0.$$

Since the polynomials $p_1(x) - 1$, $p_2(x)$, $p_3(x)$ and $p_4(x) - 1$ have integral coefficients uniquely determined by G and since λ_1, λ_2 are roots of a polynomial $(x - a)^p - 2^{2p+1}$, which, by Lemma 3, are irreducible over the rational field, it follows that

$$p_1(\lambda_2) - 1 = p_2(\lambda_2) = p_3(\lambda_2) = p_4(\lambda_2) - 1 = 0,$$

and consequently that $\lambda_2 > 4$ is not free. This is a contradiction by Theorem 2. Thus λ_1 is shown to be free. Since λ_1 is algebraic and since ϵ is an arbitrary positive number, w is a limit of free algebraic numbers. Thus, Theorem 3 is proved.

THEOREM 4. *Let a, b, c, d, k and h be non-zero integers such that $k > 2$, $(k, h) = 1$. Then*

$$\lambda = \frac{-(a+c)(b+d) \pm [(a+c)^2(b+d)^2 - 16abcd\sin^2(h\pi/k)]^{\frac{1}{2}}}{4abcd}$$

is not free.

Proof. By an elementary computation we see that the trace of the matrix $M = A^aB^bA^cB^d$ is

$$2 + 2(a+c)(b+d)\lambda + 4abcd\lambda^2 = e^{2h\pi i/k} + e^{-2h\pi i/k}.$$

Since $\det(M) = 1$, it follows that $r_1 = e^{2h\pi i/k}$ and $r_2 = e^{-2h\pi i/k}$ are characteristic roots of M. Moreover, $k > 2$, $(k,h) = 1$ implies $r_1 \neq r_2$. Therefore M can be diagonalized with diagonal elements r_1 and r_2, and hence $M^k = E$. Thus F_λ is not free.

COROLLARY 1. *Every number λ on the segment $[-2,2]$ of the real axis is a limit of non-free real numbers.*

Proof. Set $a = b = c = d = \pm 1$ in Theorem 4, and note that the numbers of the form $\cos(h\pi/k)$ are densely distributed in the segment $[0, 1]$.

COROLLARY 2. *Every number of the form λi, where $-1 < \lambda < 1$, is a limit of non-free pure imaginary numbers.*

Proof. Set $a = b = -c = -d = \pm 1$ in Theorem 4.

The authors have been unable to decide whether the domain $\{\lambda|\ |\lambda| < 1$ or $|\lambda - 1| < 1$ or $|\lambda + 1| < 1\}$ contains an open set consisting of free numbers only.

4. An example of free products. We have seen that in certain cases $F_{\alpha,\beta}$ is the free product of cyclic groups $\{A\}$ and $\{B\}$. Using similar methods we may also construct an example of free products of two abelian groups, each of which is a free abelian group of rank 2.

THEOREM 5. *Let*

$$A_1 = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \alpha i \\ 0 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 0 \\ \beta i & 1 \end{pmatrix},$$

where $|\alpha| > 2$ and $|\beta| > 2$. Then the group $F = \{A_1, A_2, B_1, B_2\}$ is the free product of two free abelian groups $F_A = \{A_1, A_2\}$ and $F_B = \{B_1, B_2\}$.

The proof of Theorem 5 is essentially the same as that of Theorem 2. The following lemma is useful.

LEMMA 4. *If*

$$z' = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}^n (z)$$

with $|\lambda| > 2$ and $|z| > 1$, then $|z'| < 1$.

Proof. Since $z' = z/(n\lambda z + 1)$, we have $z'^{-1} = n\lambda + z^{-1}$. Therefore $|z'^{-1}| > 1$ or $|z'| < 1$.

Proof of Theorem 5. Suppose F is not the free product of F_A and F_B, so that there exists an element G of F such that

$$G = B_{n_r} A_{m_r} \ldots B_{n_1} A_{m_1} = E,$$

where

$$\dot{A}_{m_i} = A_1^{m_{1i}} A_2^{m_{2i}},\ B_{n_j} = B_1^{n_{1j}} B_2^{n_{2j}}$$

We may always assume that none of

$$A_{m_i} \quad \text{and} \quad B_{n}. \qquad (1 < j < n - 1)$$

are the identity elements of F_A and F_B respectively. Then each

$$A_{m_j} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

with $|a| > 2$ and each

$$B_{n_j} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \qquad (1 < j < n - 1)$$

with $|b| > 2$. Using Lemma 4, it is easy to show, as in the proof of Theorem 2, that

$$|B_{n_{r-1}} A_{m_{r-1}} \dots B_{n_1} A_{m_1}(0)| < 1,$$
$$|A_{m_r}^{-1} B_{n_r}^{-1}(0)| > 1.$$

This gives the necessary contradiction, and Theorem 5 is proved.

Let $H_m = \{A_1{}^m B_1 A_1{}^{-m}\}$, $m = 1, 2, 3, \dots$ and $K_n = \{A_2{}^n B_1 A_2{}^{-n}, A_2{}^n B_2 A_2{}^{-n}\}$, $n = 1, 2, 3, \dots$. Each H_m is a free abelian group of rank 1 and each K_n is a free abelian group of rank 2 all contained in F. It is not hard to verify that F contains the free product of all H_m and all K_n. Therefore we can obtain a representation of the free product of any finite or countable number of free abelian groups each of which has rank 1 or 2.

The authors, however, have been unable to obtain matric representations of free products of free abelian groups whose ranks are greater than 2.

We show, as an example, that the matrices

$$C_1 = \begin{pmatrix} 1 & m & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & 0 & m \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 1 & 0 & 0 \\ m & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m & 0 & 1 \end{pmatrix}$$

do not generate a free product. The groups $F_C = \{C_1, C_2\}$ and $F_D = \{D_1, D_2\}$ are free abelian and of rank 2, and the groups $\{C_i, D_j\}$ $(i = 1, 2; j = 1, 2)$ are all free groups of rank 2. But the group $F = \{C_1, C_2, D_1, D_2\}$ is not a free product of F_C and F_D, since the relation

$$C_1C_2D_1{}^{-1}D_2C_1{}^{-1}C_2{}^{-1}D_1D_2{}^{-1}C_1C_2D_1D_2{}^{-1}C_1{}^{-1}C_2{}^{-1}D_1{}^{-1}D_2 = E$$

always holds.

References

1. J. L. Brenner, *Quelques groupes libres de matrices*, C. R. Acad. Sci. Paris *241* (1955), 1689–1691.

2. K. Goldberg and M. Newman, *Pairs of matrices of order two which generate free groups*, Ill. J. Math. *1* (1957), 446–448.

3. I. N. Sanov, *A property of a representation of a free group*, Doklady Akad. Nauk (N.S.) *57* (1947), 657–659.

4. B. L. van der Waerden, *Modern algebra*, vol. 1 (New York, 1949).

University of British Columbia

A NOTE ON DIVISION ALGORITHMS IN IMAGINARY QUADRATIC NUMBER FIELDS

D. W. DUBOIS AND A. STEGER

An integral domain E is said to be *Euclidean* if there exists a non-negative, integer-valued function g defined on the non-zero elements of E such that for every non-zero x and y in E,

(1) $g(xy) \geqslant g(x)$;

(2) (division algorithm) if x does not divide y then there exists an element q in E, depending on x and y, with

$$g(y - qx) < g(x).$$

The function g will be called a *Euclidean function.*

The elementary properties of Euclidean domains may be found in Van der Waerden (**4**, p. 56).

The problem of determining all quadratic number fields $K(\sqrt{m})$ in which the norm is a Euclidean function (on the sub-domain of algebraic integers in $K(\sqrt{m})$) has been solved. See (**2**, ch. xiv) for a partial discussion and bibliography. The following is unsolved: are there any Euclidean quadratic fields for which the norm is not a Euclidean function? That is, can the norm be generalized so as to enlarge the class of fields possessing division algorithms? The following theorem asserts that for *imaginary* quadratic fields the answer is no; the proof, based on the scarcity of units in these fields, fails for the real fields. This theorem answers a question of Hasse (**3**) concerning whether the field $K(\sqrt{-19})$, known by Dedekind (**1**, suppl. xi, p. 451) to be a principal ideal domain in which the norm is not a Euclidean function, is Euclidean in the general sense defined above, and appears to be the first proof that a principal ideal domain need not be Euclidean.

THEOREM. *An imaginary quadratic field $K(\sqrt{m})$ is Euclidean if and only if the norm N is a Euclidean function.*

Proof. The norm N is a Euclidean function for imaginary $K(\sqrt{m})$ only when $m = -1, -2, -3, -7, -11$; see (**2**) for a proof. Let $m < 0$ be different from these and suppose that $K(\sqrt{m})$ is Euclidean with Euclidean function g. There exists an integer t in $K(\sqrt{m})$ distinct from zero and units, such that $g(t)$ is a minimum of the set of all $g(x)$ for which x is neither zero nor a unit. Then for every integer b there is an integer q with $b - qt$ either zero or a unit; this means that every integer in $K(\sqrt{m})$ is congruent to zero or to a unit (mod t). But the only units are ± 1. It follows that

$$N(t) = N((t)) \leqslant 3.$$

Received May 13, 1957.

But for the m chosen above, this inequality implies that t is zero or a unit, contrary to the choice of t. The contradiction establishes the theorem.

References

1. L. Dirichlet and R. Dedekind, *Vorlesungen über Zahlentheorie* (4 Aufl. Braunschweig, 1894).
2. G. H. Hardy and E. M. Wright, *The Theory of Numbers* (Oxford, 1954).
3. Helmut Hasse, *Ueber eindeutige Zerlegung in Primelemente oder Primhauptideale in Integraetsbereichen*, J. reine angew. Math., *159* (1928), 3–12.
4. B. L. Van der Waerden, *Modern Algebra* (New York, 1949).

University of New Mexico

DIMENSION OF IDEALS IN POLYNOMIAL RINGS

MAURICE AUSLANDER AND ALEX ROSENBERG

1. Introduction. A well-known theorem asserts that if K is a field, $\mathfrak{P}$ a prime ideal in the polynomial ring $S = K[X_1, \dots X_n]$ and d the transcendence degree of $S/\mathfrak{P}$ over K

$$n = \text{rank } \mathfrak{P} + d.$$

In the first half of this paper we extend this result to the case of arbitrary commutative noetherian K, as well as giving a purely homological proof of the classical theorem. In the second half we use our first result to compute the analogue of the dimension of the product and intersection of two affine varieties when K is a Dedekind ring. This seems to be of some interest in view of (**4**).

We shall adhere to the following notations throughout: K will always be a commutative noetherian ring with unit and $S = K[X_1, \dots X_n]$ the ring of polynomials in n indeterminates over K. For a prime ideal $\mathfrak{P}$ in S with $\mathfrak{P} \cap K = \mathfrak{p}$, the *dimension* of $\mathfrak{P} = d(\mathfrak{P})$, is the transcendence degree of the field of quotients of $S/\mathfrak{P}$ over the field of quotients of $K/\mathfrak{p}$; $r(\mathfrak{P})$ is the rank of $\mathfrak{P}$. If R is a local ring, dim R is the Krull dimension = the rank of the maximal ideal = the minimal number of non-zero generators of an ideal containing some power of the maximal one. Finally, if M is a module over a ring R, $hd_R M$ is the projective dimension of M (**2**, p. 109), $w.\ hd_R M$ is the weak dimension of M (**2**, VI, Ex.3) and gl. dim R is the global dimension of R (**2**, p. 111).

2. Dimension.

LEMMA 1. *Let R be a local ring, $\mathfrak{a}$ and $\mathfrak{b}$ proper ideals in R such that $\mathfrak{b} \supset \mathfrak{a}^s$ for some integer s. If $\mathfrak{b}$ can be generated by t elements, then* dim $R \leqslant$ dim $R/\mathfrak{a} + t$.

Proof. Let $\mathfrak{m}$ be the maximal ideal in R and dim $R/\mathfrak{a} = r$. Then there exist r elements $x_1, \dots, x_r$ in $\mathfrak{m}$ such that $\mathfrak{m}^n \subset (x_1, \dots, x_r) + \mathfrak{a}$ for some integer $n > 0$. Then $\mathfrak{m}^{ns} \subset (x_1, \dots, x_r) + \mathfrak{a}^s \subset (x_1, \dots, x_r) + \mathfrak{b}$. Thus $(x_1, \dots, x_r) + \mathfrak{b}$ is an $\mathfrak{m}$-primary ideal generated by $r + t$ elements and therefore dim $R \leqslant r + t$.

THEOREM 2. *Let P be a prime ideal in the polynomial ring S, $\mathfrak{p}$ the prime ideal $K \cap \mathfrak{P}$ in K. Then*

$$d(\mathfrak{P}) + r(\mathfrak{P}) = n + r(\mathfrak{p}).$$

Proof. Let Q be the field of quotients of $K/\mathfrak{p}$. Then the natural epimorphism $K_\mathfrak{p} \rightarrow Q$ induces an epimorphism $\phi : K_\mathfrak{p}[X] \rightarrow Q[X]$ where X denotes the

Received June 7, 1957. Written with the support of the National Science Foundation.

n-tuple $(X_1, \dots, X_n)$. Let $\overline{\mathfrak{P}}$ be the prime ideal $K_{\mathfrak{p}}[X]\mathfrak{P}$ in $K_{\mathfrak{p}}[X]$. Since $\overline{\mathfrak{P}}$ contains $\mathfrak{p}K_{\mathfrak{p}}[X] = Ker\, \phi$, the ideal $\phi\,(\overline{\mathfrak{P}})$ is a prime ideal in $Q[X]$. Now $K_{\mathfrak{p}}[X]/\overline{\mathfrak{P}}$ has the same field of quotients as $K[X]/\mathfrak{P}$. Thus the transcendence degree of

$$Q[X]/\phi(\overline{\mathfrak{P}}) = K_{\mathfrak{p}}[X]/\overline{\mathfrak{P}}$$

over Q is $d(\mathfrak{P})$. Since Q is a field, we have the well-known classical result (for which we give a homological proof in Proposition 3) that

$$n = d(\mathfrak{P}) + \text{rank}\, \phi\, (\overline{\mathfrak{P}}).$$

Therefore to complete the proof it suffices to show that rank $\phi(\overline{\mathfrak{P}}) =$ rank $\mathfrak{P}$ − rank $\mathfrak{p}$.

The epimorphism $\phi : K_{\mathfrak{p}}[X] \to Q[X]$ induces an epimorphism

$$K_{\mathfrak{p}}[X]_{\overline{\mathfrak{P}}} \to Q[X]_{\phi(\overline{\mathfrak{P}})} \text{ with kernel } \quad \mathfrak{p}K_{\mathfrak{p}}[X]_{\overline{\mathfrak{P}}}\,.$$

If we let

$$R = K_{\mathfrak{p}}[X]_{\overline{\mathfrak{P}}} \text{ we have } R/\mathfrak{p}R = Q[X]_{\phi(\overline{\mathfrak{P}})}\,.$$

Using the facts that passage to rings of quotients and polynomial rings do not change the ranks of prime ideals (**5**, p. 57, p. 67) we deduce the following equalities: dim $R/\mathfrak{p}R$ = rank $\phi(\overline{\mathfrak{P}})$, rank $\mathfrak{p}$ = rank $\mathfrak{p}R$ and dim R = rank $\overline{\mathfrak{P}}$. Therefore the equality we wish to prove becomes dim R = dim $R/\mathfrak{p}R$ + rank $\mathfrak{p}R$. Let $x_1, \dots, x_r$ in $\mathfrak{p}K_{\mathfrak{p}}$ be a system of parameters in $K_{\mathfrak{p}}$. Then for some integer s, we have that $(\mathfrak{p}K_{\mathfrak{p}})^s \subset (x_1, \dots, x_r)$ and therefore $(\mathfrak{p}R)^s \subset (x_1, \dots, x_r)R$. Applying Lemma 1 we have that dim $R \leqslant$ dim $R/\mathfrak{p}R$ + rank $\mathfrak{p}R$. The reverse inequality follows trivially from the definition of dim R.

PROPOSITION 3. *Let K be a field, $\mathfrak{P}$ a prime ideal in S. Then*

$$n = d(\mathfrak{P}) + r(\mathfrak{P}).$$

Proof. Let R be the local ring $S_{\mathfrak{P}}$. Since $F = R/\mathfrak{P}R$ is isomorphic (as a K-algebra) to the field of quotients of $S/\mathfrak{P}$, the transcendence degree of F over K is $d(\mathfrak{P}) = d$. Let $t_1, \dots, t_d$ in R have the property that their images in F form a transcendence base for F over K. We consider F to be a module over the polynomial ring $R[Y_1, \dots, Y_d]$ by defining $Y_i f = t_i f$ for all $f \in F$. Then to prove Proposition 3 we shall compute $hd_{R[Y]}F$ in two different ways.

Since $K[Y] \subset R[Y]$, the set $N = K[Y] - \{0\}$ is a multiplicatively closed subset of $R[Y] - \{0\}$. Then $R[Y]_N = (R \otimes_K K[Y])_N = R \otimes_K K(Y)$ and[1] $F_N = F$. Thus, applying (**2**, VII, Ex. 10) and (**2**, VI, Ex. 3b) we have[2]

$$hd_{R[Y]}F = w.hd_{R[Y]}F = w.hd_{R[Y]_N}F_N = hd_{R \otimes K(Y)}F.$$

Since K is a field, we know that

[1] For the definition of F_N see (**2**, VII Ex. 9).

[2] The unadorned $\otimes$ refers to a tensor product over K.

$$(1)\qquad \operatorname{Ext}^q{}_{R\otimes R}(R, \operatorname{Hom}_{K(Y)}(F, C)) = \operatorname{Ext}^q{}_{R\otimes K(Y)}(F, C)$$

for all q and all $R \otimes K(Y)$-modules C (**2**, IX, 4.3). Let T be the multiplicatively closed subset of $S \otimes S$ consisting of all $h \otimes k$ where $h,k \notin \mathfrak{P}$. Consider the exact sequence

$$0 \to \mathfrak{I} \to S \otimes S \xrightarrow{\phi} S \to 0$$

where $\phi(f \otimes g) \to fg$ and $\mathfrak{I} = \operatorname{Ker} \phi$. Since $\phi(T)$ does not contain 0, the set T does not meet $\mathfrak{I}$. Therefore we obtain the exact sequence

$$0 \to \mathfrak{I}_T \to (S \otimes S)_T \to S_T \to 0.$$

It is easily seen that $(\mathrm{S} \otimes \mathrm{S})_T = R \otimes R$ and that $S_T = R$, that is, we have an exact sequence

$$0 \to \mathfrak{I}_T \to R \otimes R \to R \to 0.$$

Clearly each ideal $\mathfrak{I}^{(k)}$ generated in $S \otimes S$ by $X_i \otimes 1 - 1 \otimes X_i$ for $i = 1, \ldots k$, is a prime ideal and $\mathfrak{I}^{(n)} = \mathfrak{I}$. Thus, each of the ideals $\mathfrak{I}_T{}^{(k)}$ in $R \otimes R$ is generated by $X_i \otimes 1 - 1 \otimes X_i$ for $i = 1, \ldots k$ and is a prime ideal, and also $\mathfrak{I}_T{}^{(n)} = \mathfrak{I}_T$. Hence, the hypotheses of (**2**, VIII, 4.2) are satisfied, and we find[3] that $hd_{R\otimes R}R = n$, so that in view of (1)

$$hd_{R\otimes K(Y)}F \leqslant n.$$

Furthermore, by the discussion on p. 153 of (**2**), it also follows that

$$\operatorname{Ext}^n{}_{R\otimes R}(R, D) = D/\mathfrak{I}_T D.$$

Let $D = \operatorname{Hom}_{K(Y)}(F, F)$. Then $(r \otimes r'g)(f) = r(g(r'f))$ for all $r \otimes r' \in R \otimes R$, $g \in \operatorname{Hom}_{K(Y)}(F, F)$ and $f \in F$. Now every element in $\mathfrak{I}_T$ is a sum of elements of the form $r \otimes 1(r' \otimes 1 - 1 \otimes r')$ (**2**, p. 168). So every element in $\mathfrak{I}_T\operatorname{Hom}_{K(Y)}(F, F)$ is a sum of elements of the form $LL'g - LgL'$, where L and L' stand for the linear transformation r and r' induce on F by multiplication. Since the Y's act on F as a transcendence basis of F over K, and since F is a finitely generated extension of K, it follows that $[F:K(Y)] < \infty$. Thus, every element of D has a well-defined trace. Then since $LL' = L'L$ the trace of $LL'g - LgL' = L'(Lg) - (Lg)L'$ is zero. Therefore every element of $\mathfrak{I}_T D$ has trace zero and so $D \neq \mathfrak{I}_T D$. Consequently

$$\operatorname{Ext}^n{}_{R\otimes R}(R, \operatorname{Hom}_{K(Y)}(F, F)) \neq 0$$

and so by (1) $hd_{R[Y]}F = n$.

We now show that $hd_{R[Y]}F = d + r(\mathfrak{P})$ which will complete the proof.

[3] Since

$$hd_{R\otimes R}R \geqslant \text{gl. dim } R$$

(**2**, IX, 7.6), we have that gl. dim. $R < \infty$. Applying (**1**, Theorem 1.10), we obtain a homological proof of the well-known fact that R is a regular local ring.

Let $t_1, \ldots, t_d$ be the elements in R such that $Y_i f = t_i f$ for all $f \in F$. If we let $Z_i = Y_i - t_i$ $(i = 1, \ldots, d)$ we have that the Z_i are algebraically independent over R, $R[Z_1, \ldots, Z_d] = R[Y]$ and $Z_i F = 0$ for all i. Thus, the operation of $R[Z]$ on F is defined by the natural epimorphism $R[Z] \to R/\mathfrak{P}R = F$ which sends each Z_i to zero. Then $\mathfrak{P}'$, the kernel of this epimorphism, is the prime ideal generated by $(\mathfrak{P}R, Z_1, \ldots, Z_d)$. Since R is a regular local ring with maximal ideal $\mathfrak{P}R$, we know that R is an integral domain, $\mathfrak{P}R = (u_1, \ldots, u_r)$ with $r = r(\mathfrak{P}) = \dim R$, and each $(u_1, \ldots, u_j)$ for $j = 1, \ldots, r$ is a prime ideal of R. It follows, therefore, that $\mathfrak{P}'$ has the $r(\mathfrak{P}) + d$ generators $u_1, \ldots, u_r, Z_1, \ldots, Z_d$, and that $(u_1, \ldots, u_j)$ for $j = 1, \ldots, r$ and $(u_1, \ldots, u_r, Z_1, \ldots, Z_i)$ for $i = 1, \ldots, d$ are prime ideals in $R[Y]$. We can, therefore, apply **(2**, VIII, 4.2**)** to find the required value for $hd_{R[Y]}F$.[4]

3. Dedekind rings. We assume throughout this section (except in Lemma 6) that K is a Dedekind ring, that is, a commutative noetherian integrally closed integral domain in which every non-zero prime ideal is maximal.

THEOREM 4. *Let $\mathfrak{P}_1$, $\mathfrak{P}_2$ be prime ideals in S such that $(\mathfrak{P}_1, \mathfrak{P}_2) \neq S$, and let $\mathfrak{J}$ be the ideal generated by $\mathfrak{P}_1 \otimes 1$ and $1 \otimes \mathfrak{P}_2$ in $S \otimes S$. Then if $\mathfrak{U}$ is a minimal prime of $\mathfrak{J}$ we have*

$$d(\mathfrak{U}) = d(\mathfrak{P}_1) + d(\mathfrak{P}_2).$$

We need two lemmas before beginning the proof.

LEMMA 5. *If $\mathfrak{P}_i \cap K = 0$, $i = 1, 2$, then $\mathfrak{U} \cap K = 0$.*

Proof. We have the exact sequence

$$0 \to \mathfrak{J} \to S \otimes S \to (S/\mathfrak{P}_1) \otimes S/\mathfrak{P}_2) \to 0.$$

Since the $\mathfrak{P}_i$'s are prime ideals such that $\mathfrak{P}_i \cap K = 0$, the $S/\mathfrak{P}_i$*'s* are integral domains containing isomorphic copies of K. Thus, the $S/\mathfrak{P}_i$'s are torsion-free K-modules which also makes $(S/\mathfrak{P}_1) \otimes (S/\mathfrak{P}_2)$ a torsion-free K-module **(2**, VII, 4.5**)**. Hence, at any rate $K \cap \mathfrak{J} = 0$. Now $\mathfrak{U}/\mathfrak{J}$ is a minimal prime belonging to zero in $(S/\mathfrak{P}_1) \otimes (S/\mathfrak{P}_2)$ and so consists entirely of zero-divisors. Thus, if $\mathfrak{U} \cap K \neq 0$, then $(S/\mathfrak{P}_1) \otimes (S/\mathfrak{P}_2)$ has K-torsion, which is a contradiction.

Before stating the next lemma we recall that a ring has been called regular in **(1)** if its local ring at each non-zero prime is regular. With this definition of regularity we have

LEMMA 6. *Let K be a regular ring and $\mathfrak{P}_0$ a prime ideal in S with $\mathfrak{P}_0 \cap K = \mathfrak{p}_0$. If $f \notin \mathfrak{P}_0$ and $\mathfrak{P}_0'$ is a minimal prime of $(\mathfrak{P}_0, f)$ with $\mathfrak{P}_0' \cap K = \mathfrak{p}_0'$ we have*

$$d(\mathfrak{P}_0') = d(\mathfrak{P}_0) - 1 + r(\mathfrak{p}_0') - r(\mathfrak{p}_0).$$

[4] In **(3**, §5, Remark 1**)** it is shown by spectral sequence arguments that $hd_{R[Z]}F = d + hd_R F$. This yields an alternative proof that $hd_{R[Y]}F = d + r(\mathfrak{P})$.

Proof. We begin by showing that S is regular. For any prime ideal $\mathfrak{P}$ in S with $\mathfrak{p} = \mathfrak{P} \cap K$, we have $S_{\mathfrak{p}} = K_{\mathfrak{p}}[X]_{\overline{\mathfrak{P}}}$, where $\overline{\mathfrak{P}} = \mathfrak{P}K_{\mathfrak{p}}[X]$. By **(1**, Theorem 1.10**)**, **(3**, Theorem 6**)** and **(1**, Theorem 4.5**)**, respectively, we have

$$\begin{gathered} \text{gl. dim } K_{\mathfrak{p}} < \infty \rightarrow \text{gl. dim } K_{\mathfrak{p}}[X] < \infty \\ \rightarrow K_{\mathfrak{p}}[X] \text{ regular} \rightarrow S_{\mathfrak{p}} \text{ regular local ring.} \end{gathered}$$

Now let

$$\overline{\mathfrak{P}}_0' = \mathfrak{P}_0' S_{\mathfrak{p}_0'} \text{ and } \overline{\mathfrak{P}}_0 = \mathfrak{P}_0 S_{\mathfrak{p}_0'}.$$

Then in

$$S_{\mathfrak{P}_0'}/\overline{\mathfrak{P}}_0$$

the ideal $\overline{\mathfrak{P}}_0'/\overline{\mathfrak{P}}_0$ is a minimal prime ideal of a non-zero principal ideal, and so by the Krull Hauptidealsatz $r(\overline{\mathfrak{P}}_0'/\overline{\mathfrak{P}}_0) < 1$. But

$$S_{\mathfrak{P}_0'}/\overline{\mathfrak{P}}_0$$

is an integral domain which shows that $r(\overline{\mathfrak{P}}_0'/\overline{\mathfrak{P}}_0) = 1$, that is, there are no prime ideals between $\overline{\mathfrak{P}}_0'$ and $\overline{\mathfrak{P}}_0$. By **(1**, Proposition 2.8**)** we then find $r(\mathfrak{P}_0') = r(\overline{\mathfrak{P}}_0') = r(\overline{\mathfrak{P}}_0) + 1 = r(\mathfrak{P}_0) + 1$. An application of Theorem 2 then yields the result.

Proof of Theorem 4. Let $\mathfrak{P}_i \cap K = \mathfrak{p}_i$. Since the non-zero prime ideals of K are maximal and $(\mathfrak{P}_1, \mathfrak{P}_2) \neq S$ only three cases arise: (a) $\mathfrak{p}_1 = \mathfrak{p}_2 = \mathfrak{p} \neq 0$, (b) $\mathfrak{p}_1 = \mathfrak{p}_2 = 0$, (c) $\mathfrak{p}_1 \neq 0, \mathfrak{p}_2 = 0$.

(a) Let $\mathfrak{p}^* = \mathfrak{p}S$. Then if $\mathfrak{P}$ is any prime ideal of S with $\mathfrak{P} \cap K = \mathfrak{p}$, we have

$$d(\mathfrak{P}) = d(\mathfrak{P}/\mathfrak{p}^*) \tag{2}$$

since $S/\mathfrak{P} \approx S/\mathfrak{p}^*/\mathfrak{P}/\mathfrak{p}^*$ and $S/\mathfrak{p}^* \approx K/\mathfrak{p}[X_1, \ldots, X_n]$.

Clearly the ideal $\tilde{\mathfrak{J}} = ((\mathfrak{P}_1/\mathfrak{p}^*) \otimes 1, 1 \otimes (\mathfrak{P}_2/\mathfrak{p}^*))$ in $(S/\mathfrak{p}^*) \otimes_K (S/\mathfrak{p}^*) = (S/\mathfrak{p}^*) \otimes_{K/\mathfrak{p}} (S/\mathfrak{p}^*)$ is $\mathfrak{J}/\mathfrak{p}(S \otimes S)$, and so $\mathfrak{U}$ maps onto a minimal prime ideal $\tilde{\mathfrak{U}}$ of $\tilde{\mathfrak{J}}$. But $K/\mathfrak{p}$ is a field and so by the field case of Theorem 4 cf. for example, **(6**; 1 §4, 1**)**

$$d(\tilde{\mathfrak{U}}) = d(\mathfrak{P}_1/\mathfrak{p}^*) + d(\mathfrak{P}_2/\mathfrak{p}^*).$$

Now $\mathfrak{U} \cap K = \mathfrak{P}_i \cap K = \mathfrak{p}$ and so (2) finishes the proof.

(b) Let Q be the field of quotients of K. Then $S_{K^*} = Q[X_1, \ldots, X_n]$ and $(S \otimes S)_{K^*} = Q[X_1, \ldots, X_n] \otimes_Q Q[X_1, \ldots, X_n]$. Moreover, by Lemma 5 $\mathfrak{P}_i \cap K = \mathfrak{U} \cap K = 0$, so that the ideals $\overline{\mathfrak{P}}_i = \mathfrak{P}_i S_{K^*}$, $\overline{\mathfrak{J}} = \mathfrak{J}(S \otimes S)_{K^*}$ and $\overline{\mathfrak{U}} = \mathfrak{U}(S \otimes S)_{K^*}$ are proper ideals with $r(\overline{\mathfrak{P}}_i) = r(\mathfrak{P}_i)$ and $r(\overline{\mathfrak{U}}) = r(\mathfrak{U})$. Since $\overline{\mathfrak{J}}$ is generated by $\overline{\mathfrak{P}}_1 \otimes 1$ and $1 \otimes \overline{\mathfrak{P}}_2$ and $\overline{\mathfrak{U}}$ is still a minimal prime ideal of $\overline{\mathfrak{J}}$ the field case of Theorem 4 again applies to give

$$d(\overline{\mathfrak{U}}) = d(\overline{\mathfrak{P}}_1) + d(\overline{\mathfrak{P}}_2).$$

By Theorem 2

$$d(\mathfrak{U}) = 2n - r(\mathfrak{U}) = 2n - r(\bar{\mathfrak{U}}) = d(\bar{\mathfrak{U}})$$
$$d(\mathfrak{P}_i) = n - r(\mathfrak{P}_i) = n - r(\bar{\mathfrak{P}}_i) = d(\bar{\mathfrak{P}}_i)$$

proving Theorem 4 in this case.

(c) Here $\mathfrak{J} \supset \mathfrak{p}_1(S \otimes S)$ and so the prime ideal $\mathfrak{U} \cap (1 \otimes S) \supset (\mathfrak{P}_2, \mathfrak{p}_1)$. Hence $\mathfrak{U} \cap (1 \otimes S)$ contains some minimal prime $\mathfrak{P}_2'$ of $(\mathfrak{P}_2, \mathfrak{p}_1)$. If $\mathfrak{J}'$ is the ideal $(\mathfrak{P}_1 \otimes 1, 1 \otimes \mathfrak{P}_2')$ in $S \otimes S$ we clearly have

$$\mathfrak{U} \supset \mathfrak{J}' \supset \mathfrak{J},$$

which shows that $\mathfrak{U}$ is also a minimal prime ideal of $\mathfrak{J}'$. But then by (b)

$$d(\mathfrak{U}) = d(\mathfrak{P}_1) + d(\mathfrak{P}_2').$$

To compute $d(\mathfrak{P}_2')$ we pass to

$$K_{\mathfrak{p}_1}[X]$$

and as usual we let $\bar{\mathfrak{P}}_2'$, $\bar{\mathfrak{P}}_2$, $\bar{\mathfrak{p}}_1$ denote the extensions of $\mathfrak{P}_2'$, $\mathfrak{P}_2$, $\mathfrak{p}_1$ to this ring. Since $\mathfrak{p}_1$ is a maximal ideal, $\mathfrak{P}_2' \cap K = \mathfrak{p}_1$ also, and so $d(\mathfrak{P}_2') = d(\bar{\mathfrak{P}}_2')$. Now

$$\bar{\mathfrak{p}}_1 = \mathfrak{p}_1 K_{\mathfrak{p}_1}[X] = (\mathfrak{p}_1 K_{\mathfrak{p}_1})K_{\mathfrak{p}_1}[X].$$

But K is a Dedekind ring which means that

$$K_{\mathfrak{p}_1}$$

is integrally closed and has dimension one. Thus it is a regular local ring (**5**, Chap. 4, Theorem 8) of dimension one[5]. Therefore (**5**, Chap. 4, Proposition 7)

$$\mathfrak{p}_1 K_{\mathfrak{p}_1} = \pi K_{\mathfrak{p}_1},$$

for some

$$\pi \text{ in } K_{\mathfrak{p}_1},$$

from which it follows that

$$\bar{\mathfrak{p}}_1 = \pi K_{\mathfrak{p}_1}[X].$$

Thus $\bar{\mathfrak{P}}_2'$ is a minimal prime of $(\bar{\mathfrak{P}}_2, \pi)$ and since

$$\bar{\mathfrak{P}}_2 \cap K_{\mathfrak{p}_1} = 0$$

we know that $\pi \notin \bar{\mathfrak{P}}_2$. Lemma 6 with

$$K = K_{\mathfrak{p}_1},\ \mathfrak{P}_0 = \bar{\mathfrak{P}}_2,\ \mathfrak{P}_0' = \bar{\mathfrak{P}}_2',\ \mathfrak{p}_0 = 0,\ \mathfrak{p}_0' = \bar{\mathfrak{p}}_1$$

[5] Since

$$\text{gl. dim } K_{\mathfrak{p}_1} < \text{gl. dim } K = 1$$

the results of (**1**) could also have been used here.

then becomes available and shows

$$d(\mathfrak{P}_2') = d(\overline{\mathfrak{P}}_2') = d(\overline{\mathfrak{P}}_2) = d(\mathfrak{P}_2)$$

which completes the proof of Theorem 4.

THEOREM 5. *Let $\mathfrak{P}_1$, $\mathfrak{P}_2$ be two prime ideals in S with $(\mathfrak{P}_1, \mathfrak{P}_2) \neq S$. Then if $\mathfrak{B}$ is any minimal prime of $(\mathfrak{P}_1, \mathfrak{P}_2)$ we have*

$$d(\mathfrak{B}) > d(\mathfrak{P}_1) + d(\mathfrak{P}_2) - n.$$

Proof. Following the usual "diagonal" method we consider the natural map $\phi : S \otimes S \rightarrow S$. The pre-image of $(\mathfrak{P}_1, \mathfrak{P}_2)$ is $(\mathfrak{J}, \mathfrak{R})$ where $\mathfrak{J} = (\mathfrak{P}_1 \otimes 1, 1 \otimes \mathfrak{P}_2)$ and $\mathfrak{R} = (X_i \otimes 1 - 1 \otimes X_i)$. Clearly $\mathfrak{W}$, the pre-image of $\mathfrak{B}$, is a minimal prime ideal of $(\mathfrak{J}, \mathfrak{R})$.

Since $\mathfrak{R} \cap K = 0$ the mapping ϕ restricted to K is an isomorphism so that $\mathfrak{W} \cap K = \mathfrak{B} \cap K$. Furthermore $(S \otimes \mathrm{S})/\mathfrak{W} = S/\mathfrak{B}$ and we see once again that $d(\mathfrak{W}) = d(\mathfrak{B})$. But $\mathfrak{W} \supset \mathfrak{J}$, therefore, for some minimal prime ideal $\mathfrak{U}$ of $\mathfrak{J}$ we have $\mathfrak{W} \supset (\mathfrak{U}, \mathfrak{R}) \supset (\mathfrak{J}, \mathfrak{R})$. Hence $\mathfrak{W}$ is a minimal prime ideal of

$$(\mathfrak{U}, \mathfrak{R}) = (\mathfrak{U}, X_1 \otimes 1 - 1 \otimes X_1, \ldots, X_n \otimes 1 - 1 \otimes X_n).$$

Thus repeated application of Lemma 6 yields

$$d(\mathfrak{W}) > d(\mathfrak{U}) - n,$$

and invoking Theorem 4 we obtain

$$d(\mathfrak{B}) = d(\mathfrak{W}) > d(\mathfrak{P}_1) + d(\mathfrak{P}_2) - n.$$

REFERENCES

1. M. Auslander and D. Buchsbaum, *Homological dimension in local rings*, Trans. Amer. Math. Soc., *85* (1957), 390–405.
2. H. Cartan and S. Eilenberg, *Homological Algebra* (Princeton, 1956).
3. S. Eilenberg, A. Rosenberg, and D. Zelinsky, *On the dimension of modules and algebras* (VIII), Nagoya Math. J., *12* (1957).
4. M. Nagata, *A general theory of algebraic geometry over Dedekind domains* I, Amer. J. Math., *78* (1956), 78–116.
5. D. G. Northcott, *Ideal Theory* (Cambridge, 1953).
6. P. Samuel, *Methodes d'algebre abstraite en geometrie algebrique*, Ergeb. d. Math. (Neue Folge), *4* (1955).

Institute for Advanced Study
Northwestern University

GROUPOÏDES AUTOMORPHES PAR LE GROUPE GÉOMÉTRİQUE ET QUASİGROUPES "ENDO"

A. SADE

ARGUMENT

L'ensemble des nombres $\in Z/n$, premiers avec l'entier n, forme un groupe (le groupe géométrique) G, par rapport à la multiplication. Etant donné un ensemble de nombres réels, M, un groupoïde Q, formé d'éléments quelconques, x, est automorphe par le groupe géométrique si (i) pour tout élément $x \in Q$ et tout nombre $m \in M$, la multiplication xm est définie; (ii) l'application $(x \rightarrow xm)$ est un automorphisme de Q.

Si M devient un semi-groupe, fini ou non, et l'application $(x \rightarrow xm)$ un endomorphisme, le groupoïde Q est dit "*endo*."

La première partie expose les diverses généralisations ou restrictions de ces définitions: anneaux, clusters **(6)**, narings **(16)**, néofields **(10)**, keys **(19)**, groupes distributifs de Burstin-Mayer, en donne des illustrations et montre la corrélation de ces ensembles algébriques avec les groupoïdes automorphes par le groupe cyclique **(12)**.

La troisième partie est consacrée aux diviseurs des "endo." La quatrième concerne leur composition, leur décomposition en p-quasigroupes, leur structure. Les "endo" jouissent de propriétés élégantes qui s'abâtardissent en se transmettant aux groupoïdes qui sont seulement automorphes par le groupe géométrique. La deuxième partie, dont la lecture n'est pas indispensable à l'intelligence des autres, traite de ces derniers pour n fini et, avec une restriction sur la parité de k, des diviseurs formés par les multiples de k.

Dans la dernière section, la construction des p-quasigroupes est ramenée à celle des "endo" d'ordre premier.

I. GÉNÉRALITÉS

1. Notations et définitions. Nous utiliserons les symboles suivants: $a \Rightarrow b$, implication, a entraîne b. $a \leftrightarrows b$, a entraîne b et b entraîne a. $a \rightarrow b$, a est appliqué sur b. $\times$, $\cdot$, $*$, $\wedge$, $\ominus$, $\oplus$, $\circ$, $\odot$, signes d'opérations. Min (a, b), le plus petit des nombres a et b. $a \cong b$, a isomorphe à b. $a|b$, a divise b. $\{a\}$, groupoïde engendré par a. $\exists$, quantificateur existentiel. R, corps des nombres réels. Q, corps des fractions rationneiles. Z, anneau des entiers rationnels. Z/n, anneau des classes résiduelles modulo n.

DÉFINITION. *Un groupoïde* **(8)**, $G(\times)$ *est automorphe par le groupe géométrique s'il existe un ensemble de nombres réels, M, tel que, pour tout $m \in M$ et pour*

Reçu le 16 août 1956.

tout élément $x \in G$ (i) *la multiplication de cet élément par m soit définie* (ii) *l'application $(x \rightarrow xm)$ soit un automorphisme.*

L'ensemble de ces applications est donc un diviseur de l'automorphe $\mathbf{A}_G$ (**15**, *p.* 40) et par conséquent M est un groupe par rapport à la multiplication usuelle des nombres. Le vocable choisi tire son origine du fait que, si G est fini d'ordre n, l'ensemble M sera le groupe multiplicatif, modulo n, des $\phi(n)$ entiers inférieurs à n et premiers avec lui, auquel Cauchy (**4**, *p.* 233) a donné le nom de groupe géométrique.

Exemple I. Le groupe des translations, le groupe des homothéties dans l'espace usuel, sont automorphes par le groupe multiplicatif de **R**.

Exemple II. Le quasigroupe du $6^{\text{ème}}$ ordre $R = \{0, 1, \ldots, 5\}$ défini par les substitutions: $S_0 = (15)$, $S_1 = (05234)$, $S_2 = (0241)(35)$, $S_3 = (03)(12)(45)$, $S_4 = (0425)(13)$, $S_5 = (01432)$, où S_i détermine la translation $(x \rightarrow x \times i)$, est automorphe par $(x \rightarrow xm)$, $m = 1$ et 5.

Contre-exemple. L'ensemble $G(\wedge)$ des vecteurs libres dans l'espace à trois dimensions est un groupoïde par rapport au produit vectoriel. Mais G n'est automorphe par aucun groupe géométrique, car:

$$\mathbf{V}m \wedge \mathbf{V}'m = (\mathbf{V} \wedge \mathbf{V}')m^2$$

Exemple III. Un groupoïde peut être automorphe par une partie seulement du groupe géométrique. Ainsi le quasigroupe $Q(\times)$, (**12**, N°1), dont la loi de composition est, sur $Z/9$:

$$x \times y = 2x - y + 3, \text{ si } x - y - 2 \text{ est premier avec 3 et}$$
$$x \times y = 2x - y, \text{ dans le cas contraire,}$$

est automorphe par le sous-groupe $m = 1, 4, 7$ du groupe géométrique $\{x \rightarrow xm\}$, (m premier avec 3). Mais pour $m = 2, 5, 8$ il se projette sur un quasigroupe différent de Q. Pour $m = 3$ ou 6, l'image de Q est un quasigroupe du $3^{\text{ème}}$ ordre, qui n'est pas un diviseur de Q.

En général, si m n'est plus premier avec n, l'application $(x \rightarrow xm)$ projette le groupoïde G, d'ordre n, sur un système qui n'est plus un groupoïde. Ainsi, dans le quasigroupe R défini ci-dessus (II), $(x \rightarrow 2x)$ n'est pas un endomorphisme et projette R sur un ensemble algébrique qui ne satisfait plus à la loi d'unicité du produit: $(2 \times 2 = [2, 0, 4])$.

2. Groupoïdes "endo." *Un groupoïde $G(\times)$ admet les endomorphismes du semi-groupe* (**18**) *des homothéties, ou plus brièvement, est "endo," s'il existe un ensemble de nombres réels, M, tel que, pour tout $m \in M$ et tous $x, y \in G$,*

(i) *le produit xm soit défini,*

(ii) *l'application $(x \rightarrow xm)$ soit un endomorphisme:*

$$x \times y = z \Rightarrow (xm) \times (ym) = zm.$$

Si G est fini, d'ordre n, les homothéties $(x \rightarrow xm)$, $(m = [0, 1, \ldots, n - 1])$ forment un semi-groupe isomorphe au semi-groupe multiplicatif de Z/n; il contient le groupe géométrique: $(m, n) = 1$.

Exemple I. Le groupe additif de Z/n est projeté par l'homothétie $(x \rightarrow xm)$, $(m, n) = k$, $n = kd$, sur son diviseur, le groupe cyclique d'ordre d:

$$(0, k, 2k, \ldots, n - k).$$

Exemple II. Soit G le groupe des translations engendré dans le plan par deux vecteurs non parallèles, $\mathbf{U}$ et $\mathbf{V}$, et dont les éléments sont de la forme:

$$\mathbf{W} = a\mathbf{U} + b\mathbf{V},$$

où a et b décrivent l'ensemble Z des entiers rationnels. Les éléments dont les coefficients a et b sont multiples d'un entier donné d, forment un diviseur D de G. Le groupe G se projette sur son diviseur D par l'endomorphisme $(x \rightarrow xd)$, car:

$$\mathbf{W} = a\mathbf{U} + b\mathbf{V} \rightleftarrows \mathbf{W}d = a\mathbf{U}d + b\mathbf{V}d.$$

Exemple III. On peut donner une définition plus générale encore et considérer des *ensembles munis de deux lois de composition* ($\times$ *et* $*$), *telles que la seconde soit distributive à droite, à gauche, ou des deux côtés par rapport à la première:* $z * (x \times y) = (z * x) \times (z * y)$.

(a) Tel est l'ensemble dont les éléments sont les sous-ensembles d'un ensemble donné, si les lois de composition sont l'inter-section $\cap$ et la réunion $\cup$.

(b) Soit G l'ensemble des polynomes de degré $q - 1$, $P = \Sigma a_i x^i$, $(i = 0, 1, \ldots, q - 1)$, où p et q sont deux entiers naturels fixes et où les coefficients a_i sont définis sur Z/p (ce qui revient, pour $x = p$, à écrire les nombres $0, 1, \ldots, p^q - 1$ dans le système de base p). On définit un groupe Abélien $G(\times)$, d'ordre p^q, en prenant pour loi de composition $P \times P' = \Sigma\ (a_i + a_i')x^i$ (cela revient, si $x = p$, à l'addition sans retenues). Si m est un entier quelconque, appelons produit de P par m le polynome $P{*}m = \Sigma(a_i m)x^i$ où le coefficient ma_i est toujours calculé modulo p. Il est évident que l'application $P \rightarrow P{*}m$ est un endomorphisme de G. Si $x = p$, quand P et m décrivent l'ensemble $0, 1, \ldots, p^q - 1$, celui-ci est muni de deux opérations, la seconde étant distributive par rapport à la première.

(c) Tel est encore l'anneau Z, si la première loi est: $x \times y = 2x - y$ et la seconde: $x{*}y = x + y$.

(d) Si la première opération définit un groupe, on obtient un "*cluster*" **(6)**. Quand ce groupe est Abélien, le cluster devient un "*naring*" **(16)**. Si de plus la seconde opération est un semi-groupe, le naring est un anneau.

Tandis que le cluster est construit en se donnant le groupe $(\times)$ et en déterminant la seconde opération $(*)$ de manière qu'elle soit distributive par rapport à la première, dans le présent travail on se donne la multiplication $(*)$ et on construit la première loi $(\times)$ à partir de la seconde.

(e) Si la première opération définit un loop **(2)** et si les éléments non nuls forment un groupe par rapport à la seconde loi, on a affaire à un "*néofield*" **(10)**.

(f) Enfin, si les deux lois coïncident, on aboutit à la loi III des groupoïdes introduits sous le nom de "kēis" par Takasaki **(19)** et connus aussi sous celui de distributifs dans le cas des quasigroupes **(3)**. La première loi (c) en offre une illustration.

3. Définitions plus restreintes. Dans ce qui suit, nous nous en tiendrons aux définitions N° 1 et 2 restreintes à des ensembles de nombres de la façon suivante:

Un groupoïde G ($\times$), défini sur Z/n, est automorphe par le groupe géométrique si, pour tout entier m, premier avec n, et pour tous $x,y,z \in G$, on a:

$$x \times y = z \Rightarrow (xm) \times (ym) = zm,$$

le produit xm ayant la signification usuelle sur Z/n.

Un groupoïde défini sur un ensemble de nombres réels, G, est "endo" si la relation précédente a lieu quels que soient $x,y,m \in G$.

4. Connexion avec Q_c. *Si n admet une racine primitive, r, tout groupoïde $Q_g(\times)$ d'ordre n, automorphe par le groupe géométrique, est (partiellement) isomorphe à un groupoïde $Q_c(*)$, automorphe par le groupe cyclique* **(12)**.

Preuve. Soit r une racine primitive de n. Les résidus (mod n) des puissances de r seront, à l'ordre près, les $\phi(n)$ nombres plus petits que n et premiers avec n. L'application $(r^i \rightarrow i)$ définit donc un système algébrique, $Q_c(*)$, automorphe par le groupe cyclique, et d'ordre $\phi(n)$. En effet, par hypothèse

$$x \times y = z \Rightarrow (xm) \times (ym) = zm, \qquad (\text{mod } n);$$

par suite (ind x) $*$ (ind y) $=$ ind z entraîne

$$(\text{ind } m + \text{ind } x) * (\text{ind } m + \text{ind } y) = \text{ind } m + \text{ind } z,$$

c'est à dire, en changeant de notation

$$a * b = c \Rightarrow (a + h) * (b + h) = c + h \qquad (\text{mod } \phi(n)).$$

Mais l'isomorphisme n'est que partiel en ce sens que Q_c est incomplet **(12, N°8)**, le produit sur Q_c n'étant pas partout défini. De plus, si tout élément de Q_c a une préimage, tout élément de Q_g n'a pas une image dans Q_c.

Si r n'était plus racine primitive, l'ordre de Q_c serait un diviseur de $\phi(n)$ et $Q_c * Q_c$ pourrait même devenir vide.

Exemple I. Prenons pour Q_g le groupe additif de $Z/9$ et 2 pour racine primitive. Aux exposants [1, 2, 3, 4, 5, 6] correspondent les restes [2, 4, 8, 7, 5, 1]. Si l'on remplace partout, dans Q_g, [2, 4, 8, 7, 5, 1] par [1, 2, 3, 4, 5, 0] on obtient un quasigroupe incomplet du 6^{ème} ordre, où le produit n'est défini que si les facteurs ont même parité, et qui est automorphe par le groupe cyclique. Sa loi de composition est $x*y = 2x - y + 1 \pmod 6$, $x \equiv y \pmod 2$.

Si l'on prend $r = 4$, les exposants sont [1, 2, 3], les restes [4, 7, 1] et, en remplacant [4, 7, 1] par [1, 2, 0] dans Q_e, on obtient un groupoïde incomplet du $3^{\text{ème}}$ ordre, où aucun produit n'est défini.

Exemple II. Soit le quasigroupe "endo" du $7^{\text{ème}}$ ordre Q_e, défini par ses translations $S_i = (x \to x \times i)$; $S_0 = (132645)$, $S_1 = (034156)$, $S_2 = (061235)$, $S_3 = (025314)$, $S_4 = (052463)$, $S_5 = (016542)$, $S_6 = (043621)$; en prenant 3 comme racine primitive de 7, les exposants sont [1, 2, 3, 4, 5, 6] et les restes [3, 2, 6, 4, 5, 1]. Supprimant l'ancien zéro et projetant chaque reste sur son indice, c.-à-d. [1, 2, 3, 4, 5, 6] sur [0, 2, 1, 4, 5, 3], on fait de Q_e un quasigroupe incomplet, du $6^{\text{ème}}$ ordre, automorphe par le groupe cyclique et défini (**12**, N° 4) par

$$S_0 = \begin{pmatrix} 012345 \\ 542\text{–}03 \end{pmatrix}.$$

Exemple III. Inversement, si n a une racine primitive, r, à tout groupoïde d'ordre $\phi(n)$, automorphe par le groupe cyclique, on pourra faire correspondre, par projection des indices sur les nombres, un groupoïde d'ordre n, automorphe par le groupe géométrique.

Ainsi, en prenant 2 comme racine primitive de 5, le quasigroupe incomplet Q, automorphe par le groupe cyclique, défini sur $Z/4$ par

$$S_0 = \begin{pmatrix} 0123 \\ 0\text{–}32 \end{pmatrix}$$

devient par la substitution

$$\begin{pmatrix} 0123 \\ 1243 \end{pmatrix}$$

un groupoïde dans lequel il suffit de remplacer les produits non définis par zéro, et de compléter en satisfaisant à la loi de cancellation, pour parvenir à un quasigroupe du $5^{\text{ème}}$ ordre, automorphe par le groupe géométrique, et ayant pour loi de composition $x \times y = 4x + 2y$, sur $Z/5$.

5. Non identité entre les propriétés des Q_c et des Q_e. La proposition 4 semble dénier tout intérêt à l'étude des groupoïdes "endo," Q_e. En réalité, presqu'aucune des propriétés des groupoïdes automorphes par le groupe cyclique ne se transmet aux "endo," car (i) ce parallélisme n'existe que si n a une racine primitive; (ii) Q_c est toujours incomplet et seulement d'ordre $(p-1)p^q$, où p est premier impair (Si $n = 4$, Q_c est vide); (iii) alors que tout entier naturel, n, est de $n - 1$ manières la somme de deux autres, il n'est égal au produit de deux facteurs que dans la mesure où il est composé.

En fait, le domaine des Q_e est plus riche que celui des Q_c de tout l'apport fourni par la notion de divisibilité.

II. GROUPOIDES AUTOMORPHES PAR LE GROUPE GÉOMÉTRIQUE

6. Lemme. Tout groupoïde "endo" est automorphe par le groupe géométrique; mais la réciproque n'est pas vraie. Les "endo" jouissent de propriétés qui ne se retrouvent pas dans les groupoïdes automorphes par le groupe géométrique et non "endo." Nous allons d'abord étudier les propriétés appartenant à ces derniers.

Si $n = sk$, la suite $j, j + s, j + 2s, \ldots, j + (k - 1)s$, où j est premier avec s, contient $\phi(n)/\phi(s)$ termes premiers avec n. Ce nombre est toujours supérieur à un sauf si s est impair et k égal à deux.

Preuve. Soit $\bar{k}$ le plus grand diviseur de k qui soit premier avec s. En supposant k et s décomposés en produits de puissances de facteurs premiers inégaux, on voit facilement que la condition nécessaire et suffisante pour que x soit premier avec n est que x soit premier avec s et $\bar{k}$. La condition $(s,\bar{k}) = 1$ implique que les résidus de $i, i + s, i + 2s, \ldots, i + (k - 1)s$ forment $k/\bar{k}$ systèmes complets de restes modulo $\bar{k}$. Ainsi, chaque telle séquence a le même nombre de termes premiers avec $\bar{k}$. Pour que ces nombres soient premiers avec n il faut et il suffit que leur premier terme i soit premier avec s. Par suite les $\phi(s)$ séquences pour lesquelles $(i,s) = 1$ contiennent chacune le même nombre de termes premiers avec n, et en même temps elles contiennent la totalité des termes premiers avec n. D'où la relation de l'énoncé.

On vérifie sans peine qu'un nombre n, et son diviseur s, ont le même indicateur dans le seul cas où s est impair et moitié de n.

7. Relation entre les éléments. *Dans un groupoïde $G(\times)$, d'ordre n, automorphe par le groupe géométrique, si les PGCD de n avec deux éléments quelconques, a et a', sont $d = (n,a)$ et $d' = (n,a')$ et si $(d,d') = k$, alors les produits $a \times a'$ et $a' \times a$ sont multiples de k en général, et de $k/2$ si k est pair et n/k impair.*

Preuve. Soit $a \times a' = a''$ et p un nombre inférieur à n et premier avec n. En vertu de l'automorphisme $(ap \times (a'p) = a''p$. Mais

$$(\exists p)\ ap \equiv a,\ a'p \equiv a' \qquad (\text{mod } n). \tag{1}$$

En effet ces deux congruences s'écrivent

$$p \equiv 1 \qquad (\text{mod } n/d),\ p \equiv 1 \qquad (\text{mod } n/d').$$

Si $s = [n/d, n/d']$, les deux conditions simultanées équivalent à

$$p \equiv 1 \qquad (\text{mod } s). \tag{2}$$

Mais, en vertu d'une relation connue $n/s = (d,d')$, on aura $n = ks$ et il existera, d'après le lemme, $\phi(n)/\phi(s)$ valeurs de p satisfaisant à (2) et premières avec n.

Pour que le produit $a \times a'$ soit univoque il faut que, pour ces valeurs de p, on ait $a''p \equiv a''$ (mod n) ou

$$a''(1 + sx) \equiv a'' \pmod{n}$$

et en divisant par s

$$a''x \equiv 0 \pmod{k}. \tag{3}$$

Soit $k = \prod A^{\alpha}$ (A premier). Si toutes les valeurs de x qui rendent $1 + sx$ premier avec n étaient divisibles par A, en posant $x = Ay$, la suite $1 + sx = 1+sAy$ contiendrait $\phi(n)/\phi(sA)$ termes premiers avec n au lieu de $\phi(n)/\phi(s)$. Mais l'égalité $\phi(sA) = \phi(s)$ n'est possible que si $A = 1$ ou 2. Si $A \neq 2$, il y a au moins une valeur de x qui n'est pas divisible par A. Pour cette valeur de x, la condition (3) entraîne $a'' \equiv 0 \pmod{A^{\alpha}}$.

Ainsi a'' est divisible par A^{α} et, le même raisonnement pouvant être fait pour tous les facteurs premiers impairs de k, si k est impair, a'' est divisible par k.

Si k est pair, le raisonnement subsiste pour tous les facteurs de k, sauf pour $A = 2$ et $\phi(sA) = \phi(s)$, ce qui suppose s impair. Soit alors $k = 2^{\alpha}k'$, $n = 2^{\alpha}k's$, k' et s impairs. Considérons de nouveau la suite $1, 1 + s, \ldots, 1 + (k - 1)s$; les valeurs impaires de x fournissent des termes $1 + sx$ pairs. Donc les termes de cette suite qui sont premiers avec n sont tout contenus dans la suivante: $1, 1 + 2s, 1 + 4s, \ldots, 1 + 2ys, \ldots, 1 + (k-2)s$, et leur nombre est $\phi(n)/\phi(2s) = \phi(n)/\phi(s)$.

Parmi les valeurs de y qui rendent $1 + 2ys$ premier avec n, il y en a au moins une qui est première avec 2 car, si toutes ces valeurs étaient paires, les termes $1 + xs$ premiers avec n seraient tous compris dans la suite $1 + 4ts$ ou, en posant $4s = s'$, dans la suite $1 + ts'$. Or cette dernière contient seulement $\phi(n)/\phi(4s) = \phi(n)/2\phi(s)$ termes premiers avec n, puisque s est impair; ce nombre n'est pas égal à $\phi(n)/\phi(s)$. Soit donc $y = 2z + 1$ la (ou une) valeur impaire de y qui rend $1 + xs$ premier avec n; on a

$$x = 2y = 2(2z + 1)$$

et la condition (3) devient

$$2(2z + 1)a'' \equiv 0 \pmod{2^{\alpha}k'}$$

ou

$$a'' \equiv 0 \pmod{2^{\alpha-1}};$$

ainsi a'' est divisible par $k/2$.

8. Autre relation. *Dans un quasigroupe $Q(\times)$, d'ordre n, automorphe par le groupe géométrique, si $a \times a' = a''$, (a, a' impairs $\in Q$), alors les PGCD, $(a,n) = d$, $(a',n) = d'$, $(a'',n) = d''$ satisfont à la relation $(d,d') = (d',d'') = (d'',d)$ et l'on a $d'' = (d,d')h$, où les facteurs premiers de n qui divisent h sont premiers avec d et d', ou bien figurent dans les décompositions, de d et de d', avec le même exposant. La proposition est vraie quels que soient a et a' si n est impair.*

Preuve. A cause de la loi du quotient, deux des congruences

$$pa \equiv a, pa' \equiv a', pa'' \equiv a'' \qquad (\mathrm{mod}\ n), \qquad (p, n) = 1,$$

doivent entraîner la troisième. D'après le N° précédent, a'' est divisible par (d,d'), autrement dit, (d,d') divise a'' et n, donc aussi (a'',n) ou d''; ainsi

$$(d, d')|d'', \qquad (d', d'')|d, \qquad (d'', d)|d',$$

et le *PGCD* de d, d', d'' est à la fois (d, d'), (d', d'') et (d'', d). On aura donc d$'' = h(d, d')$.

Désignons par A les facteurs premiers de n et soit

$$d = \prod A^{\alpha}, d' = \prod A^{\alpha'}, d'' = \prod A^{\alpha''},$$

on aura

$$(d, d') = \prod A^{\mathrm{Min}(\alpha,\alpha')}$$
$$(d', d'') = \prod A^{\mathrm{Min}(\alpha',\alpha'')}$$
$$(d'', d) = \prod A^{\mathrm{Min}(\alpha'',\alpha)}.$$

Donc, pour tout facteur premier de n,

$$\mathrm{Min}(\alpha,\alpha') = \mathrm{Min}(\alpha',\alpha'') = \mathrm{Min}(\alpha''\alpha).$$

En vertu de l'égalité $d'' = h(d,d')$ le facteur A ne peut diviser h que si $\alpha'' > \mathrm{Min}(\alpha,\alpha')$, ce qui exige $\alpha = \alpha'$.

9. Diviseurs. *Si $Q(\times)$ est un quasigroupe d'ordre n, automorphe par le groupe géométrique, et k un diviseur impair de n, les multiples de k forment un sous-quasigroupe D, de Q, d'ordre n/k, isomorphe par $(x \to x/k)$ à un quasigroupe $E(.)$ qui est lui-même automorphe par le groupe géométrique d'ordre $\phi(n/k)$.*

Preuve. Soit $(n,a) = d$ et $(n,a') = d'$. Alors

$$k|n,\ k|a \text{ et } k|a' \Rightarrow k|(d,d'),$$

donc (N° 7) $a \times a'$ est multiple de (d,d') et par conséquent de k. Ainsi l'ensemble des multiples de k est fermé et forme un sous-quasigroupe de Q. Si dans celui-ci on divise tous les éléments par k, on trouve un quasigroupe, $E(.)$, d'ordre n/k, composé des éléments $[0, 1, 2, \ldots, n/k - 1]$ et où nous définissons la composition $(\cdot)$ par

$$a \times a' = a'' \Rightarrow (a/k).(a'/k) = a''/k$$

donc $D \cong E$. Mais, à cause de l'automorphisme

$$a \times a' = a'' \Rightarrow (ap) \times (a'p) = a''p, \qquad (p,n) = 1, p < n.$$

Donc $(a/k).(a'/k) = a''/k \Rightarrow (ap/k).(a'p/k) = a''p/k$.

L'ensemble des valeurs de p parcourt (mod n/k) tous les nombres premiers avec n/k. En effet, d'après le lemme, si l'on remplace les $\phi(n)$ nombres plus petits que n et premiers avec $n = ks$ par leurs résidus (mod s), on obtient les $\phi(s)$ nombres plus petits que s et premiers avec s, chacun le même nombre de fois. En faisant $n/k = s$, on voit que la condition

$$(a/k).(a'/k) = a''/k \Rightarrow (ap/k).(a'p/k) = a''p/k \qquad (\text{mod } n/k)$$

est vérifiée pour toute valeur de p première avec n/k; ainsi le quasigroupe $E(.)$ est bien automorphe par le groupe géométrique (x → px), $(p,n/k) = 1$.

Quand Q est "endo" la proposition est vraie même sans la restriction sur la parité de k, et la preuve est immédiate. Mais elle n'est pas valable si Q n'est pas "endo."

10. Corrélation avec Q_e. Avant de passer a l'étude des "endo," observons que, en vertu du N° 4, les propositions 10 et suivantes de **(12)** ont ici leurs corrélatives:

Si Q $(\times)$ *est un quasigroupe d'ordre n automorphe par le groupe géométrique,* $(x \times y = z)$, *son conjoint* $(y.x = z)$, *son réciproque* $(z \ominus y = x)$, *le quasigroupe* R $(\odot)$*défini par* $x \odot y = az$, $(a,n) = 1$, *le quasigroupe* S $(*)$*défini par* $ax{*}ay = z$ *et le quasigroupe* T $(\wedge)$, *image de* Q *par la transformation* $x \rightarrow x^a$, $(a,\phi(n)) = 1$, *avec la définition* $x^a \wedge y^a = z^a$, *sont encore automorphes par le groupe géométrique.*

La vérification est immédiate pour les quatre premiers cas. Dans le dernier, observons que le groupe géométrique est invariant par $(x \rightarrow x^a)$ si a est premier avec l'ordre $\phi(n)$ de ce groupe. Alors, si $x \times y = z$ et par suite $xm \times ym = zm$, $(m,n) = 1$, *on aura*

$$x^a \wedge y^a = z^a \Rightarrow (xm)^a \wedge (ym)^a = (zm)^a \Rightarrow m^a x^a \wedge m^a y^a = m^a z^a$$

et comme m^a décrit les mêmes valeurs (mod n) que m,

$$mx^a \wedge my^a = mz^a.$$

Exemple. Soit A le quasigroupe du $9^{\text{ème}}$ ordre [0, 1, 2, . . . , 8], automorphe par le groupe géométrique et défini par $1 \times 0 = 1$, $1 \times 3 = 2$, $2 \times 3 = 1$, $A \times 1 = [8, 0, 6, 4, 5, 3, 7, 1, 2]$, et où le diviseur $D = [0, 3, 6]$ a pour loi de composition, sur $Z/9$, $x \times y = x + 2y$. Si l'on fait subir à A l'isomorphisme (2, 5)(4, 7), qui laisse invariant le groupe géométrique d'ordre $\phi(9)$, on obtient un nouveau quasigroupe qui est encore automorphe par le groupe géométrique. (Il suffit de le vérifier en prenant pour m une racine primitive $\rho = 2$.)

Plus généralement, *si* $Q(\times)$ *est un quasigroupe automorphe par le groupe géométrique* G, *et si* R (.) *est isomorphe à* Q *par* T: $(x \rightarrow x')$, *une condition suffisante pour que* R *soit automorphe par* G *est que* T *laisse* G *invariant.*

Car $mx \times my = mz \leftrightarrows (mx)'.(my)' = (mz)'$, mais puisque $T \in A_G$, $(mx)' = m'x' = \mu x'$, donc $\mu x'.\mu y' = \mu z'$, *avec* $\mu \in G$.

11. Tables. Pour $n = 1, 2, 3$ on obtient seulement la solution ordinaire

$$x \times y = ax + by \pmod{n}, \qquad (a, b \text{ premiers avec } n)$$

qui est toujours "endo." Toutefois, pour $n = 2$, on a la solution exceptionnelle: $x \times y = x + y + 1$.

Pour $n = 4$ on trouve le groupe carré de *K*lein, son produit par l'isotopie **(1)**

$$\xi = 1, \qquad \eta = 0.13.2, \qquad \zeta = 1,$$

et le conjoint de ce dernier.

Pour $n = 5$, une prospection exhaustive donne 32 solutions, 16 ordinaires et 16 définies par les produits $Q \times a = [1, 0, 2, 4, 3], [1, 2, 3, 0, 4], [1, 3, 4, 2, 0], [1, 4, 0, 3, 2]$, $(a = 1, 2, 3, 4)$, en complétant de manière à respecter la loi du quotient.

12. Isomérie. *Si, dans un quasigroupe $Q(\times)$, d'ordre n, automorphe par le groupe géométrique, $(x \rightarrow xm)$, $(m,n) = 1$, on remplace par isomérie* **(11**, pp. 8-10**)** *l'ensemble produit $D \times D$ induit par Q sur le sous-quasigroupe D de Q composé des multiples de k, $(k|n)$, par l'ensemble produit $D'*D'$, où D' est un quasigroupe automorphe par $(x \rightarrow xm)$, et composé des mêmes éléments que D, alors on obtient un nouveau quasigroupe Q' qui est encore automorphe par le groupe géométrique.*

Ainsi, si Q et R sont deux quasigroupes du même ordre n, automorphes par le groupe géométrique, on obtiendra deux nouveaux quasigroupes, automorphes par le groupe géométrique, en échangeant par isomérie le diviseur $D = 0,d,2d, \ldots, n-d$, avec $(n,d) = d$ du premier contre le diviseur $D' = 0,d,2d, \ldots$ du second.

D'un quasigroupe "endo" on pourra déduire par isomérie une série de quasigroupes automorphes par le groupe géométrique; toutefois toutes les solutions ne seront pas atteintes de cette façon, car on peut construire des quasigroupes automorphes par le groupe géométrique et dont le diviseur $D = 0,d,2d, \ldots$ n'est pas normal, (N° 14). Ainsi, dans l'exemple du N° 10, le diviseur $D = [0, 3, 6]$ n'est pas normal.

Néanmoins la remarque précédente, d'une part met en lumière l'importance de la notion d'*isomérie*, introduite en 1950 **(11)** et que l'on rencontre d'une façon toute naturelle et presque obligatoire dans l'étude des quasigroupes, dont elle est une propriété caractéristique; d'autre part, elle nous mène à l'étude des quasigroupes "endo."

III. QUASIGROUPES "ENDO"

13. "Endo" usuels. Avant d'aborder l'étude des "endo," signalons quelques quasigroupes susceptibles de servir d'illustration aux propriétés que nous rencontrerons plus loin.

(a) Sur Z/n la loi

$$x \times y = ax + by \qquad (a, b, \text{ premiers avec } n)$$

définit un quasigroupe "endo."

(i) L'équation à gauche $x \times c = d$, ou $ax + bc = d$, a une solution unique $x = (d-bc)a'$, où a' est l'associé de a, $(aa' = 1)$. Il en est de même pour l'équation à droite $c \times y = d$. Donc $A(\times)$ est un quasigroupe.

(ii) A est "endo" car

$$xm \times ym = amx + bmy = m(ax + by) = (x \times y)m.$$

(b) Sur le corps Q des fractions rationnelles, la loi

$$x \times y = ax + by, \qquad (a, b \neq 0 \text{ dans } Q)$$

définit un quasigroupe "endo," car

(i) $$(\exists x)\, x \times c = d$$

puisque $ax + bc = d$ a une solution unique $x = (d-bc)\,/a$, a n'étant pas nul; et de même $(\exists y)\; c \times y = d$. Donc B est un quasigroupe.

(ii) On voit comme pour A que B est "endo."

(c) Le quasigroupe B reste évidemment "endo" si l'on suppose a et b entiers.

(d) Sur le corps R des nombres réels, la loi

$$x \times y = ax + by, \; (a, b \text{ quelconques } \neq 0)$$

définit un quasigroupe "endo." Si a et b sont rationnels, B est un diviseur de D.

(e) Si a et b sont entiers, D est un "endo" dont C est diviseur.

14. Diviseur. (i) *Si a est un nombre réel quelconque, convenons d'appeler multiples de a les produits de a par les entiers rationnels* $Z = [0, \pm 1, \pm 2, \ldots]$. *Deux nombres sont congrus par rapport à a si leur différence est multiple de a.* (ii) *La condition nécessaire et suffisante pour que, dans un quasigroupe "endo,"* $Q(\times)$, *les multiples d'un élément quelconque, a, forment un diviseur D_a est que les éléments entiers de Q forment eux-mêmes un sous-quasigroupe de Q. Si Q est fini, d'ordre n,* $D_a = [0, d, 2d, \ldots, n-d]$, $d = (n,a)$ *et D_a est d'ordre n/d.* (iii) *Si* $Q(\times)$ *est construit dans le corps R des nombres réels, ou dans celui, Q, des fractions rationnelles tout nombre* $a \neq 0$ *dans* $Q(\times)$ *définit une partition des éléments de Q où deux éléments appartiennent ou non au même bloc suivant qu'ils sont, ou non, congrus par rapport à a. La condition nécessaire et suffisante pour que cette partition soit régulière est que la partition définie par* $a = 1$ *le soit. D_a est alors normal dans Q. Le système des représentants est l'ensemble des éléments de Q qui ne sont pas congrus les uns des autres par rapport à a. Le quasigroupe quotient Q/D_a a même puissance que l'ensemble des éléments de Q compris entre* 0 *et a.*

Preuve. (ii) La condition est nécessaire car, en faisant $a = 1$, les multiples de 1, c'est-à-dire les éléments entiers de Q, forment un diviseur.

Elle est suffisante car si

$$x, y \in Z \Rightarrow x \times y \in Z,$$

soient ax et ay deux multiples de a. Leur produit sera, en tenant compte de l'endomorphisme,

$$ax \times ay = a(x \times y),$$

ce qui est un multiple de a puisque $x \times y$ est entier par hypothèse. Donc le complexe des multiples de a est fermé parrapport à (×). Les équations $q \times x = b$ et $x \times q = b$ ont une solution unique dans ce complexe, donc (**8**,p. 986) celui-ci est un quasigroupe.

Si Q est fini d'ordre n, le diviseur existera toujours et sera $D = [0,d,2d, \ldots ,n-d]$ (mod n), d'ordre n/d, car les multiples de a s'indentifient avec ceux de $d = (a,n)$.

Exemple I. Dans le quasigroupe A (N° 13) défini par

$$x \times y = 2x + 8y \qquad \text{(mod 15)},$$

les quatre diviseurs sont [0], [0, 5, 10], [0, 3, 6, 9, 12] et $[A]$ lui-même.

Exemple II. Dans le groupe additif des fractions rationnelles (C, N° 13, avec $a = b = 1$), le complexe des entiers relatifs forme un diviseur, et il est isomorphe au diviseur $[0, \pm p/q, \pm 2p/q, \ldots]$, formé par les multiples de la fraction fixe arbitraire $p/q \neq 0$.

(iii) Si i et i' sont deux éléments de Q, non congrus par rapport à a, ils représentent deux classes disjointes

$$\Sigma\, az + i \quad \text{et} \quad \Sigma\, az + i', \quad z \in Z,$$

de la partition déterminée par a sur Q. Le produit (×) de deux éléments $az + i$ et $az' + i'$ peut se mettre sous la forme $az'' + i''$, où i'' est bien déterminé, à une congruence près par rapport à a:

$$(1) \qquad (az + i) \times (az' + i') = az'' + i''.$$

Si cette partition est régulière, i'' ne doit dépendre que de i et de i' (par rapport à a). En vertu de l'endomorphisme, si l'on multiplie les trois éléments de (1) par $1/a$, on aura;

$$(2) \qquad (z + \ i/a) \times (z' + i'/a) = z'' + i''/a.$$

Mais si deux nombres i et j sont congrus (ou non) par rapport à a, les quotients i/a et j/a seront en même temps congrus (ou non) par rapport à 1 et vice-versa, car

$$a|(i-j), \text{ ou } i-j = ka \leftrightarrows i/a - j/a = k,$$

où k est entier, c'est-à-dire multiple de 1.

Les égalités (1) et (2) expriment que les partitions définies sur Q par a et par 1 sont régulières en même temps.

Si D_1 est le diviseur de Q formé de ses éléments entiers et D_a celui qui est formé par les multiples de a, D_1 et D_a sont alors normaux. D'ailleurs ils sont isomorphes car, à cause de l'endomorphisme

$$D_1 \cong D_a = D_1(x \to xa)$$

Le quasigroupe quotient Q/D_a est formé des cosets $az + i$, $az + i'$, . . . , $z \in Z$, $(i-i')/a$ non entier.

On peut ramener tous les i entre 0 et a; le système des représentants est formé de tous les éléments de Q compris entre 0 et a.

Exemple. (N° 13,C), $x \times y = ax + by$. La fraction 3/4 définit une partition régulière dont le diviseur normal est formé des multiples de 3/4 et dont les cosets sont $\Sigma\, 3z/4 + p/q$, où p/q n'est pas multiple de 3/4; on a

$$(3z/4 + p/q) \times (3z'/4 + p'/q') = 3(az + bz')/4 + ap/q + bp'/q'.$$

Le représentant $ap/q + ap'/q'$ ne dépend que des représentants p/q et p'/q'. A chaque fraction comprise entre 0 et 3/4 correspond un coset. Le quasigroupe quotient est isomorphe au quasigroupe défini par la même loi que C, sur l'intervalle (0, 3/4), modulo 3/4.

15. Diviseur Normal. *Si $Q(\times)$ est un quasigroupe "endo" d'ordre n, si f est un élément quelconque de Q et si $(f,n) = d$, $(n = kd)$,* (i) *l'endomorphisme $(x \to xf)$ projette Q sur son diviseur $D =$* [0, d, $2d$, . . . , $n-d$]. (ii) *Celui-ci est toujours normal, les cosets sont $\Sigma\, ud + i$, $(u = 0, 1, \ldots, k-1)$.* (iii) *Le quasigroupe quotient Q/D est isomorphe au diviseur* [0, k, $2k$, . . . , $n-k$]. (iv) *Si E est un ensemble quelconque d'éléments $\in Q$, en nombre k, respectivement congrus* (mod k) *à* [0, 1, 2, . . . , $k-1$]. *l'ensemble produit $E.E = E \times E$* (mod k) *est isomorphe à D par $(x \to xd)$.* (v) *Enfin E (.) est "endo"* (mod k).

Preuve. (i) Dans l'application T $(x \to xf)$, tous les éléments x, $x + k$, $x + 2k$, . . . , $x + n-k$ ont la même image xf. Ainsi, T projette homomorphiquement Q sur celui de ses diviseurs, D, qui est composé des multiples de f, c'est-à-dire de d.

(ii) La partition modulo d est régulière car, quels que soient u et v, on a par $(x \to xk)$

$$du + i \to ki \quad (i < d)$$
$$dv + j \to kj \quad (j < d)$$
$$(ud + i) \times (vd + j) \to ki \times kj = k(i \times j).$$

Mais les éléments qui se projettent sur $k(i \times j)$ sont

$$i \times j,\ i \times j + d,\ i \times j + 2d, \ldots, i \times j + n - d,$$

donc $(ud + i) \times (vd + j) = wd + (i \times j)$;
par suite les cosets $\Sigma_u\, ud + i$ et $\Sigma_v\, vd + j$ ont pour produit (×) le coset $\Sigma_W\, wd + (i \times j)$.

(iii) Le système des représentants est $0, 1, 2, \ldots, i, \ldots, j, \ldots, d-1$. Le quasigroupe quotient Q/D est isomorphe au diviseur $0, k, 2k, \ldots, n-k$, puisque

$$i \times j = r \leftrightarrows (ki) \times (kj) = kr.$$

(iv) Soit E un système quelconque de k éléments respectivement congrus à $0, 1, 2, \ldots, k-1 \pmod k$. L'application $T = (x \rightarrow xf)$ les projette sur

$$0, f, 2f, 3f, \ldots, n-f \qquad (\text{car } kf = 0, \bmod n)$$

c'est-à-dire univoquement sur les éléments de D.

Soient $x, y \in E$ et $x \times y = z$. En tenant compte de l'endomorphisme de Q, on a

$$(fx) \times (fy) = fz \qquad \pmod n$$

Mais cette égalité a lieu aussi modulo n/d. L'ensemble produit $E \times E$ est composé d'éléments qui, ramenés au dessous du module k, forment un quasigroupe $E(.)$ isomorphe à D par T.

(v) Reste à montrer que $E(.)$ est "endo" (mod k). Par hypothèse, Q admet les endomorphismes $(x \rightarrow xm)$, $(m < n)$. Toute application $(x \rightarrow xm)$ projette D sur un de ses diviseurs, pouvant coïncider avec D. Cela signifie que

$$(ad) \times (bd) = cd \Rightarrow (mad) \times (mbd) = mcd \pmod n.$$

Mais puisque $D(\times) \cong E(.)$ par T,

$$a.b = c \Rightarrow (ma).(mb) = mc \qquad \pmod k.$$

Exemple. Soit le quasigroupe "endo" du $15^{\text{ème}}$ ordre $Q = [0, 1, 2, \ldots, 14]$ défini par les produits:

$$Q \times 1 = [11, 4, 0, 8, 7, 6, 14, 10, 3, 2, 1, 9, 5, 13, 12];$$
$$1 \times Q = [2, 4, 6, 5, 13, 12, 14, 1, 0, 8, 7, 9, 11, 10, 3].$$

Prenons $f = 10$. Par $(x \rightarrow 10x)$ on le projette sur son diviseur $D_5 = [0, 5, 10]$.

Pour $d = 3$, on a la partition $D_3 = [0, 3, 6, 9, 12]$; $C = [1, 4, 7, 10, 13]$; $C' = [2, 5, 8, 11, 14]$. Le quasigroupe quotient Q/D_3 est isomorphe à $D_5 = [0, 5, 10]$.

Pour $d = 5$, on a le diviseur normal $D_5 = [0, 5, 10]$ et les cosets $C = [1, 6, 11]$; $C' = [2, 7, 12]$; $C'' = [3, 8, 13]$; $C''' = [4, 9, 14]$. Le quasigroupe quotient est isomorphe à D_3.

Si l'on prend $E = [2, 3, 5, 11, 14] \equiv [2, 3, 0, 1, 4] \pmod 5$ et si l'on remplace tous les éléments de E et de $E \times E$ par leur reste (mod 5) on obtient $E(.) \cong D_3 \cong Q/D_5$.

IV. COMPOSITION DES "ENDO"

16. Lemme. Si deux quasigroupes K et S, d'ordres k et s, sont "endo", leur produit direct **(9)** sera encore un quasigroupe Q et, d'après ce que l'on sait du produit direct, en regardant l'endomorphisme (x → xm) comme une opération externe, distributive, Q admettra encore cette opération distributive. Mais pareille affirmation a une signification illusoire, car les éléments de Q ne sont plus des nombres, mais des êtres complexes (x,y), $x \in K$, $y \in S$. Les démonstrations suivantes sont donc nécessaires. Nous utiliserons le lemme connu **(7**, Théorème 59**)**:

Si $n = sk$, $(s, k) = 1$, *tout nombre* (mod n) *peut se mettre, d'une manière et une seule, sous la forme* $kx + sy$, $x < s$, $y < k$.

17. Produit direct. *Le produit direct de deux "endo" est "endo,"* ou plus précisément: *Si* $(s,k) = 1$ *et si* $K\ (\times) = [0,1, \dots ,y, \dots ,k-1]$ *et* $S(.) = [0, 1, \dots , x ,\dots , s-1]$ *sont deux "endo" (ou deux quasigroupes automorphes par le groupe géométrique), le groupoïde* $G(*)$*défini par*

$$G = \sum (kx + sy),\ (xk + sy) * (kx' + sy') \equiv k(x.x') + s(y \times y'), \quad (\text{mod } ks),$$

est un quasigroupe "endo" (ou un quasigroupe automorphe par le groupe géométrique).

Preuve. A tout couple ordonné $kx + sy$, $kx' + sy'$ correspond un produit et un seul; donc G est un groupoïde.

Supposons que

$$(kx + sy)*(kx' + sy') \equiv (xk + sy)*(kx'' + sy'') \qquad (\text{mod } n)$$

donc $x.x' - x.x'' \equiv 0$ (mod s) et $x' = x''$; pareillement $y' = y''$. Le calcul est le même à droite. Ainsi $Q(*)$ est un quasigroupe.

Pour que Q soit "endo" (automorphe par le groupe géométrique) il faut qu'il satisfasse, pour tout facteur f (mod n), (pour tout facteur f premier avec n) à

$$[f(kx + sy)]*[f(kx' + sy')] \equiv fk(x.x') + fs(y \times y') \qquad (\text{mod } n).$$

Si l'on cherche les coefficients de u et v, de k et de s, lorsqu'on met le nombre $f(kx + sy)$ sous la forme $ku + sv$ (mod n), $u < s$, $v < k$, on trouve que u et v sont uniques et bien déterminés. Si l'on appelle u' et v' les quantités analogues à u et v, relativement à x' et y', la condition devient

$$(ku + sv)*(ku' + sv') \equiv fk(x.x') + fs(y \times y') \qquad (\text{mod } n),$$

ou finalement

$$k(u.u') + s(v \times v') \equiv k(u.u') + s(v \times v') \qquad (\text{mod } n),$$

où $u.u'$ est défini (mod s) et $v \times v'$ (mod k), ce qui est une identité. Dans le cas des "endo" elle est vérifiée quel que soit f. Si les quasigroupes sont automorphes par le groupe géométrique, elle est satisfaite pour tout nombre f, premier avec k et s, c'est-à-dire avec n.

Exemple I.

$$k = 2\ , s = 3;\ K(\times) = 0, 1, x \times y = x + y \qquad (\text{mod } 2);$$

$$S(.) = [0, 1, 2];\ x.y = 2(x + y) \qquad (\text{mod } 3);$$

$$Q(*) = [0, 1, 2, 3, 4, 5];\ x*y = 5(x + y) \qquad (\text{mod } 6).$$

Exemple II. $k = 5$; $s = 3$. Prenons pour $K(\times)$ le quasigroupe défini par $S_0 = (1\,2\,4\,3)$, $S_1 = (0\,1\,4\,2)$, $S_2 = (0\,2\,3\,4)$, $S_3 = (0\,3\,2\,1)$, $S_4 = (0\,4\,1\,3)$, où S_i est la translation $(x \rightarrow x \times i)$. Comme quasigroupe $S(.)$, gardons le même $x.y = 2(x + y)$ (mod 3) que dans l'exemple précédent.

Alors $Q(*)$ a pour loi de composition

$$(5x + 3y)*(5x' + 3y') \equiv 5(x.x') + 3(y \times y') \qquad (\text{mod } 15).$$

Il coïncide avec le quasigroupe donné en exemple au N° 15.

Les composants S et K ne sont autre chose que les quasigroupes $E(.)$ définis au N° 15 et relatifs aux modules s et k (*cf.* l'exemple du N° 15).

18. Réciproque. *Tout "endo" d'ordre ks, $(k, s) = 1$, est le produit direct de deux "endo" d'ordres k et s;* ou, plus précisément: *Si Q $(*)$ est un quasigroupe "endo," d'ordre $n = sk$, $(k, s) = 1$, alors $\exists$ deux "endo" $K(\times)$ d'ordre k et S $(.)$ d'ordre s, respectivement isomorphes à $D_s = [0, s, 2s, \ldots, n-s]$ par $(x \rightarrow xs)$ et à $D_k = [0, k, 2k, \ldots, n-k]$ par $(x \rightarrow xk)$, tels que, pour tous $a = kx + sy$, $a' = kx' + sy'$; $x, x' < s$; $y, y' < k$, on ait*

$$a*a' \equiv k(x.x') + s(y \times y') \qquad (\text{mod n}).$$

Preuve. Soit $Q(*)$ un quasigroupe "endo" d'ordre $n = ks$, $(k, s) = 1$. Si a et a' sont deux éléments quelconques de Q, chacun d'eux peut, d'une seule manière, être mis sous la forme

$$a = kx + sy,\ a' = kx' + sy'$$

avec $x, x' < s$; $y, y' < k$; et $a, a' \in Q$.

On a vu (N° 15, (ii)) que si

$$a \equiv i, \quad a' \equiv i' \qquad (\text{mod } s),$$

alors

$$a*a' \equiv i*i' \qquad (\text{mod } s),$$

or on a

$$a \equiv kx, \quad a' \equiv kx' \qquad (\text{mod } s).$$

Donc

$$a*a' \equiv (kx)*(kx') \equiv k(x*x') \qquad (\text{mod } s).$$

Considérons le quasigroupe $S(.) = [0, 1, 2, \dots, s-1]$, induit par $(*)$ de la manière suivante. Soit l'ensemble produit $E*E$, où $E = [0, 1, 2, \dots, s-1]$. Remplaçons tous les éléments de $E*E$ par leur reste (mod s). Nous obtenons un quasigroupe $S(.)$ qui (N° 15, (iv) et (v)) est "endo" et isomorphe au diviseur $D_k = [0, k, 2k, \dots, n-k]$ par $(x \rightarrow xk)$. Ainsi on a

$$x*x' \equiv x.x' \qquad (\text{mod } s),$$

d'où

$$a*a' \equiv k(x.x') \qquad (\text{mod } s).$$

On trouve pareillement

$$a*a' \equiv s(y \times y') \qquad (\text{mod k}),$$

où $K(\times)$ est induit par $(*)$ sur l'ensemble $F*F$, $(F = [0, 1, 2, \dots, k-1])$, pris modulo k, et est isomorphe au diviseur $D_s = [0, s, 2s, \dots, n-s]$ par $(x \rightarrow xs)$.

On a donc

$$a*a' \equiv k(x.x') + s(y \times y') \qquad (\text{mod } s \text{ et mod } k)$$

et comme s et k sont premiers entre eux, la congruence est vérifiée modulo n.

La réciproque n'est pas vraie, en général, pour les quasigroupes automorphes par le groupe géométrique (voir N° 1, exemple II).

19. Lemme. Citons la proposition connue **(5, II**, p. 64**) (17**, p. 130**)**.

Si $n = a^\alpha b^\beta c^\gamma \dots$, *(a,b,c, ... premiers inégaux), tout entier* A (mod n) *peut être décomposé d'une manière et une seule en une somme*

$$A = (n/a^\alpha)x + (n/b^\beta)y + (n/c^\gamma)z + \dots, 0 \leqslant x < a^\alpha, 0 \leqslant y < b^\beta, \dots.$$

20. Généralisation. *Tout "endo" fini est le produit direct d'"endo" ayant pour ordres des puissances de nombres premiers,* ou plus précisément: *Si* $n = a^\alpha b^\beta c^\gamma \dots$ *(a,b,c ... premiers inégaux), tout quasigroupe "endo,"* $Q(*)$, *d'ordre* n, *a pour loi de composition*

$$A * A' = \sum_a (n/a^\alpha)(x \times x')_a, \text{ où } A = \sum_a (n/a^\alpha)x \in Q,$$

$A' = \Sigma_a (n/a^\alpha)x' \in Q; 0 \leqslant x,x' < a^\alpha$ *et où* $K(\times)_a$ *est le quasigroupe "endo" d'ordre* a^α, *induit par* $x*x' \equiv (x \times x')_a$ (mod a^α), *isomorphe au diviseur de* Q: $[0, (n/a^\alpha), (2n/a^\alpha), \dots, (n - n/a^\alpha)]$ *par* $(x \rightarrow (nx)/a^\alpha)$.

Preuve. On sait (N° 15, (ii)) que

$$A \equiv i, A' \equiv i' \quad (\text{mod } s), \qquad (s,n) \equiv s \Rightarrow A*A' \equiv i*i' \qquad (\text{mod } s).$$

Or

$$A \equiv (n/a^\alpha)x, \qquad A' \equiv (n/a^\alpha)x' \qquad (\text{mod } a^\alpha),$$

donc

$$A*A' \equiv (nx/a^\alpha)*(nx'/a^\alpha) \qquad (\text{mod } a^\alpha).$$

A cause de l'endomorphisme

$$(nx/a^\alpha)*(nx'/a^\alpha) \equiv (n/a^\alpha)(x*x') \qquad (\text{mod } n).$$

Donc

$$A*A' \equiv (n/a^\alpha)(x \times x')_a \qquad (\text{mod } a^\alpha).$$

Considérons la somme

$$\sum{}_a(n/a^\alpha)(x \times x')_a.$$

Tous ses termes, sauf le premier, ont des coefficients divisibles par a^α, et comme le premier est congru à $A*A'$ (mod a^α), on a

$$A * A' \equiv \sum{}_a(n/a^\alpha)(x \times x')_a \qquad (\text{mod } a^\alpha).$$

La même chose peut être répétée pour toutes les puissances $b^\beta, c^\gamma, \ldots$. Finalement

$$A * A' \equiv \sum{}_a(n/a^\alpha)(x \times x')_a \qquad (\text{mod } a^\alpha, b^\beta, c^\gamma, \ldots),$$

donc aussi (mod n).

21. Semi-groupes des "endo". *L'ensemble des quasigroupes "endo" finis est, par rapport à l'opération de composition des "endo"* (N° 17) *un semi-groupe incomplet et commutatif, homomorphe à celui des entiers naturels, où la multiplication* (*usuelle*) *n'est supposée être définie que si les facteurs sont premiers entre eux, l'image de tout "endo" d'ordre n étant le nombre n.*

Preuve. L'opération de composition entre deux "endo" dont les ordres k et s sont premiers entre eux, définie au N° 17, est associative et commutative; car, si k, s, t sont trois nombres premiers entre eux deux à deux, et si $n = kst$, soient K,S,T trois "endo" d'ordres respectifs k, s, t. Si l'on compose K et S, et le résultat avec T, ou S et T, puis K avec ce produit, on obtiendra un "endo" d'ordre n, dont la loi de composition, d'après le N° 20, sera dans les deux cas

$$A * A' \equiv \sum{}_a(n/a^\alpha)(x \times x')_a \qquad (\text{mod } n).$$

On peut d'ailleurs le vérifier par un calcul facile. Soient $K(*)$, $S(\times)$ *et* $T(.)$ les trois "endo" d'ordres respectifs k,s,t. Composons K et S; la loi de multiplication du résultat sera $(\oplus)$.

$$X \oplus X' \equiv (sx + ky) \oplus (sx' + ky') \equiv s(x*x') + k(y \times y') \qquad (\text{mod } ks).$$

Composons ce quasigroupe avec T; le quasigroupe résultant sera $Q(\circ)$

$$A \circ A' = (tX + ksz) \circ (tX' + ksz') \equiv t(X \oplus X') + ks(z.z') \qquad (\text{mod } n)$$

$$ou \quad A \circ A' = st(x*x') + tk(y \times y') + ks(z.z') \qquad (\text{mod } n).$$

Cette expression est indépendante de l'ordre des trois composants K, S, T. Le produit est donc associatif.

Ainsi, si l'on considère l'ensemble de tous les "endo" d'ordre fini, l'opération de composition organise cet ensemble en un semi-groupe commutatif incomplet, où la composition n'est définie que si les deux "endo" composants ont des ordres premiers entre eux. L'"endo" unité est le quasigroupe du premier ordre $0 \times 0 = 0$. Un "endo" n'a pas d'inverse; mais la loi d'existence du quotient est satisfaite toutes les fois que l'ordre de l'"endo" dividende, n, est un multiple de l'ordre, k, de l'"endo" diviseur, avec $n = ks$ et $(k,s) = 1$. Quand le quotient existe il est unique. Ce semi-groupe incomplet est homomorphe au semi-groupe multiplicatif des entiers naturels, dans lequel la multiplication est supposée n'être définie que si les deux facteurs sont premiers entre eux. Tous les "endo" d'un ordre déterminé, n, se projettent sur l'élément n.

22. Lemme. *Si $d = a^{\alpha'}b^{\beta'}c^{\gamma'} \ldots$ est un diviseur de $n = a^{\alpha}b^{\beta}c^{\gamma} \ldots$ (a, b, c . . . premiers inégaux), alors tout multiple de d, inférieur à n, a pour décomposition suivant le* N° 19

$$B = (n/a^{\alpha})a^{\alpha'}u + (n/b^{\beta})b^{\beta'}v + \ldots \pmod{n}$$

où u est défini modulo

$$a^{\alpha-\alpha'},$$

v modulo

$$b^{\beta-\beta'}, \ldots ;$$

et tout nombre ayant une telle décomposition est multiple de d.

On vérifie aisément que B est multiple de toutes les

$$a^{\alpha'}.$$

23. Diviseur engendré par deux autres. *Le sous-quasigroupe engendré dans un "endo" $Q(*)$, d'ordre n, par les diviseurs $D_d = [0, d, 2d, \ldots, n-d]$ et $D_{d'} = [0, d', 2d', \ldots n-d']$, $(d|n, d'|n)$, est le diviseur $D_{d''} = [0, d'', 2d'', \ldots n-d'']$, où $d'' = (d, d')$.*

Preuve. Soit

$$d' = a^{\alpha''}b^{\beta''}c^{\gamma''} \ldots$$

et P le produit, dans $Q(*)$, d'un multiple de d par un multiple de d',

$$P = [\sum (n/a^{\alpha})a^{\alpha'}u) * (\sum (n/a^{\alpha})a^{\alpha''}u'] \equiv \sum (n/a^{\alpha})(a^{\alpha'}u \times a^{\alpha''}u')_a;$$

$$P = \sum (n/a^{\alpha})a^{\min(\alpha',\alpha'')}(\omega \times \omega')_a \qquad \pmod{n},$$

où

$$\omega = u, \omega' = a^{\alpha''-\alpha'}u' \text{ si } \alpha'' > \alpha'$$

et

$$\omega' = u', \omega = a^{\alpha'-\alpha''}u \text{ si } \alpha'' < \alpha'.$$

Cela montre que P est multiple de

$$\prod a^{\min(\alpha',\alpha'')} = d'', \qquad (\text{N}^\circ 7)$$

De plus, u décrit toutes les valeurs

$$0, 1, \ldots, a^{\alpha-\alpha'} - 1,$$

et u' toutes les valeurs

$$0, 1, \ldots, a^{\alpha-\alpha''} - 1.$$

Si $\alpha'' \geqslant \alpha'$, ω décrit toutes les valeurs de u et, dans le quasigroupe $(\times)_a$, le produit $(\omega \times \omega')_a$ les parcourt aussi.

Si $\alpha'' \leqslant \alpha'$, ω' décrit les

$$a^{\alpha-\alpha'}$$

valeurs de u' et il en est de même de $(\omega \times \omega')_a$. Donc, dans les deux cas, ce produit prend

$$a^{\alpha-\min(\alpha',\alpha'')}$$

valeurs distinctes. Le nombre des valeurs parcourues par l'élément P, c'est-à-dire l'ordre de

$$\{D_d, D_{d'}\}$$

est

$$\prod \frac{a^\alpha}{a^{\min(\alpha',\alpha'')}} = \frac{n}{d''} \cdot$$

Ce nombre est celui des multiples de d'' dans Q. Ainsi

$$\{D_d, D_{d'}\} = D_{d''}.$$

24. Diviseurs admissibles. Treillis. (i) *Tout diviseur admissible d'un quasigroupe "endo," $Q(*)$, d'ordre n, est composé des multiples d'un entier d, $(d|n)$. On peut dire que ces sous-quasigroupes admissibles sont les idéaux* (**16**, p. 94) *de l'"endo."* (ii) *Sur un quasigroupe "endo" d'ordre n, le treillis* (lattice) *des sous-quasigroupes D_d, $(d|n)$, est isomorphe au treillis formé par les sous-groupes du groupe cyclique C_n.*

Preuve. (i) Soit $D = [a, b, c, \ldots]$ un diviseur admissible (au sens de "zulässig", (**14**)). Les endomorphismes de Q projettent D sur l'ensemble des éléments $am, bm, cm, \ldots$ quel que soit $m \in Q$. Donc D contient, en même temps que a, tous les nombres $[a, 2a, 3a, \ldots, (n-1)a, 0]$, autrement dit $Dx \subseteq D$ pour tout $x \in Q$.

Soit $(a,n) = d$. Cette suite contient seulement n/d termes distincts (mod n), à savoir $[0, d, 2d, \ldots, n-d]$, à l'ordre près. Donc D contient le diviseur D_d formé par les multiples de d.

Si cet ensemble n'épuise pas tous les éléments de D, soit $b \in D$ et non multiple de d. Désignons par d' le $PGCD$ de b et de n, $(b,n) = d'$. Alors D

contiendra à la fois tous les multiples de d et, pour la même raison, tous les multiples de d', et par conséquent le sous-quasigroupe engendré par les multiples de d et de d', lequel est formé (N° 23) par l'ensemble des multiples de (d,d'). En répétant ce raisonnement jusqu'à ce que tous les éléments de D soient épuisés, on voit que D est le sous-quasigroupe composé des multiples d'un diviseur de n. (ii) Il est clair, d'autre part, que

$$D_d \cap D_{d'} = D_\Delta,$$

où Δ est le $PPCM$ de d et d'.

25. Décomposition en p-quasigroupes. *Dans tout quasigroupe $Q(*)$ "endo" d'ordre n $(n = a^\alpha b^\beta c^\gamma \dots)$, le sous-quasigroupe $\{A\}$, engendré par l'élément $A = \Sigma_a(n/a^\alpha)x$, est le produit direct des diviseurs engendrés par les éléments x, respectivement dans chaque quasigroupe $(x \times x')_a$, d'ordre a^α.*

Preuve. Pour construire $\{A\}$ il faut former les puissances de A, puis les produits de ces puissances deux à deux, et ainsi de suite, jusqu'à fermeture. Or

$$A * A = (n/a^\alpha)(x \times x)_a + (n/b^\beta)(y \times y)_b + \dots.$$

Pendant ces opérations successives, l'élément x, dans le quasigroupe "endo" $(x \times x')_a$ d'ordre a^α, engendre un diviseur d'ordre a'. De même $\{y\}$, dans $(y \times y')_b$ est un diviseur d'ordre b', etc. Donc $\{A\}$ sera le produit direct, d'ordre $a'b'c' \dots$, de ces divers quasigroupes; symboliquement

$$\{A\} = \Sigma(n/a^\alpha)\,\{x\}.$$

Ainsi les problèmes de construire un "endo" d'ordre donné et de trouver ses diviseurs monogènes, sont ramenés aux problèmes analogues pour les quasigroupes "endo" dont l'ordre est une puissance d'un nombre premier, ou *p-quasigroupes.*

V. P-QUASIGROUPES "ENDO"

26. Isotopie. Soit $Q(\times)$ un "endo" d'ordre $n = p^\alpha$ où p est premier. On a vu (N° 15) que, pour toute valeur β $(0 \leqslant \beta < \alpha)$ le diviseur $D_\beta = [0, p^\beta, 2p^\beta, \dots, n-p^\beta]$ est normal; les cosets sont

$$C_{\beta,i} = [i, i + p^\beta, i + 2p^\beta, \dots, i + n-p^\beta] \qquad (i = [0, 1, \dots, p^\beta - 1]).$$

et l'on a

$$C_{\beta,i} \times C_{\beta,j} = C_{\beta,i\times j}.$$

Le quasigroupe quotient est isomorphe au diviseur $D_{\alpha-\beta}$ et aussi à l'ensemble produit $R_\beta \times R_\beta(\text{mod } p^\beta)$, où R_β est l'ensemble des restes $[0, 1, \dots, p^\beta - 1]$. (On le voit en intervertissant k et d dans le N° 15 (iv).) Enfin, $D_{\alpha-\beta} = [0, p^{\alpha-\beta}, 2p^{\alpha-\beta}, \dots, (p^\beta - 1)p^{\alpha-\beta}]$ se projette par $(x \rightarrow x/p^{\alpha-\beta})$ sur un quasigroupe d'ordre p^β, "endo."

Exemple. Soit le quasigroupe "endo" $Q = [0, 1, 2, \ldots, 8]$ défini par $1 \times Q = [2, 1, 3, 5, 7, 6, 8, 4, 0]$; $Q \times 1 = [5, 1, 3, 8, 7, 6, 2, 4, 0]$. $D_1 = [0, 3, 6]$. L'ensemble produit $[0, 1, 2] \times [0, 1, 2]$ est isomorphe (mod 3) à D_1: $R \times R = D_1(3x \to x)$. En posant $C_0 = [0, 3, 6]$; $C_1 = [1, 4, 7]$; $C_2 = [2, 5, 8]$, le quasigroupe quotient est $(C_0, C_1, C_2) \cong D_1$.

THÉORÈME. *Si $Q(\times)$ est un quasigroupe "endo" d'ordre $n = p^\alpha$, (p premier) (i) par l'application $(x \to k)$, où k est le quotient entier $[x/p^\beta]$ de x par p^β, l'ensemble produit des cosets $C_{\beta,i} \times C_{\beta,j} = C_{\beta,r}$, où $r = i \times j$ et $i, j, r \neq 0$, $C_{\beta,i} = [i,\ i + p^\beta,\ i + 2p^\beta, \ldots, i + kp^\beta, \ldots, i + n - p^\beta]$, se projette sur un quasigroupe $G_{i,j}(.)$ d'ordre $p^{\alpha-\beta}$. Si $\alpha \geqslant 2\beta$, G est invariant par l'autotopie* **(1)**

$$A\left\{\xi = \begin{pmatrix} x \\ x + hi \end{pmatrix}, \eta = \begin{pmatrix} y \\ y + hj \end{pmatrix}, \zeta = \begin{pmatrix} z \\ z + hr \end{pmatrix}\right\}$$

où $h = p^{\alpha-2\beta}$.

(ii) *Si $\beta = \alpha - 1$ et $D = [0, p^{\alpha-1}, 2p^{\alpha-1}, \ldots, kp^{\alpha-1}, \ldots, (p-1)p^{\alpha-1}]$, le quasigroupe quotient Q/D est isomorphe à $D_1 = [0, p, 2p, \ldots, n-p]$ et se projette par $(x \to x/p)$ sur un quasigroupe "endo" d'ordre $p^{\alpha-1}$. Si $i, j, r \neq 0$, $G_{i,j}(.)$ est isotope par*

$$T\left\{\xi = \begin{pmatrix} x \\ xi' \end{pmatrix}, \eta = \begin{pmatrix} y \\ yj' \end{pmatrix}, \zeta = \begin{pmatrix} z \\ zr' \end{pmatrix}\right\}$$

$$ii' \equiv jj' \equiv rr' \equiv 1 \qquad \pmod p$$

à un quasigroupe automorphe par le groupe cyclique **(12)**

(iii) *Si $j = 0$, et $i,r \neq 0$, $G_{i,0}(.)$ est isotope par*

$$T'\left\{\xi' = \begin{pmatrix} x \\ xi' \end{pmatrix}, \eta' = \begin{pmatrix} y \\ 0.y \end{pmatrix}\zeta', \zeta' = \begin{pmatrix} z \\ zr' \end{pmatrix}\right\}$$

au groupe cyclique $x \wedge y = x + y$ (mod p). *On a un résultat symétrique si $i = 0$, au lieu de j.*

(iv) *Si $r = 0$ et $i,j \neq 0$, $G_{i,j}(.)$ est isotope au même groupe cyclique par*

$$T''\left\{\xi'' = \begin{pmatrix} x \\ xi' \end{pmatrix}, \eta'' = \begin{pmatrix} y \\ -yj' \end{pmatrix}, \zeta'' = \begin{pmatrix} z.0 \\ zi' \end{pmatrix}\right\}$$

(v) *Si $r = j = i = 0$, $G_{0,0}(.) \cong D$ par $(x \to x/p^{\alpha-1})$*

Preuve. (i) Considérons l'ensemble produit

$$C_{\beta,i} \times C_{\beta,j} = C_{\beta,r}, \qquad r = i \times j,$$

dans la partition régulière définie par D_β. On a

$$(i + kp^\beta) \times (j + k'p^\beta) \equiv r + k''p^\beta \qquad (\text{mod } p^\beta).$$

Appliquons l'endomorphisme et multiplions par $(p^{\alpha-\beta} + 1)$:

$$(i + kp^\beta)(p^{\alpha-\beta} + 1) \equiv i + ip^{\alpha-\beta} + kp^\beta \equiv i + (k + ip^{\alpha-2\beta})p^\beta \qquad (\text{mod } p^\alpha)$$

$$[i + (k + ip^{\alpha-2\beta})p^\beta] \times [j + (k' + jp^{\alpha-2\beta})p^\beta) = r + (k'' + rp^{\alpha-2\beta})p^\beta.$$

Ainsi le quasigroupe $G_{i,j}(.)$ obtenu en retranchant i à tous les éléments du multiplicande, j à tous ceux du multiplicateur, r à tous les produits, puis en divisant les restes par p^{β}, $G = [0, 1, 2, \ldots, p^{\alpha-\beta} - 1]$, satisfait à la relation

$$k.k' = k'' \Rightarrow (k + ih).(k' + jh) = k'' + rh \qquad (\operatorname{mod} p^{\alpha-\beta})$$

où $h = p^{\alpha-2\beta}$. Il coïncide avec lui-même par une isotopie dont les trois composantes sont des substitutions régulières (**4**, p. 162.):

$$A\left\{\xi = \begin{pmatrix} x \\ x + hi \end{pmatrix} ; \eta = \begin{pmatrix} y \\ y + hj \end{pmatrix} ; \zeta = \begin{pmatrix} z \\ z + hr \end{pmatrix}\right\} \qquad (\operatorname{mod} p^{\alpha-\beta})$$

En particulier, si $i = j = r$, l'autotopie devient un automorphisme par un diviseur du groupe cyclique.

(ii) Soit $\beta = \alpha - 1$ et appelons D le diviseur d'ordre p

$$D = [0, p^{\alpha-1}, 2p^{\alpha-1}, 3p^{\alpha-1}, \ldots, kp^{\alpha-1}, \ldots, (p - 1)p^{\alpha-1}].$$

Il définit une partition régulière; les cosets sont

$$C_{\alpha-1,i} = [i, i + p^{\alpha-1}, i + 2p^{\alpha-1}, \ldots, i + p^{\alpha-1}k, \ldots, i + (p - 1)p^{\alpha-1}].$$

Considérons l'ensemble produit

$$(1) \qquad C_{\alpha-1,i} \times C_{\alpha-1,j} = C_{\alpha-1,r}; \quad i \times j = r \qquad (\operatorname{mod} p^{\alpha-1}).$$

On a

$$(kp^{\alpha-1} + i) \times (k'p^{\alpha-1} + j) = k''p^{\alpha-1} + r \qquad (\operatorname{mod.} p^{\alpha})$$

Multiplions les trois éléments par $1 + p^{\alpha-1}$

$$(kp^{\alpha-1} + i)\ (1 + p^{\alpha-1}) \equiv i + p^{\alpha-1}(k + i) \qquad (\operatorname{mod} p^{\alpha})$$

car $p^{2\alpha-2} \equiv 0$ (mod p^{α}) puisque $\alpha > 1$.

Donc, en appliquant l'endomorphisme

$$(2) \qquad [p^{\alpha-1}\ (k + i) + i] \times [p^{\alpha-1}\ (k' + j) + j] \equiv p^{\alpha-1}\ (k'' + r) + r \ (\operatorname{mod} p^{\alpha})$$

Formons le quasigroupe $G(.)$ d'ordre p, ($G = [0, 1, 2, \ldots, p-1]$) déduit de $C_i \times C_j$ par $x \rightarrow [x/p^{\alpha-1}]$

$$G = (C_i \times C_j)\begin{pmatrix} x \\ k \end{pmatrix};$$

d'après (2), G jouit de la propriété

$$(3) \qquad k.k' = k'' \Rightarrow (k + i).(k' + j) = k'' + r \qquad (\operatorname{mod} p)$$

Il coïncide avec lui-même par l'isotopie (**1**) dont les composantes sont trois substitutions circulaires

$$A\left\{\xi = \begin{pmatrix} x \\ x + i \end{pmatrix} ; \eta = \begin{pmatrix} y \\ y + j \end{pmatrix} ; \zeta = \begin{pmatrix} z \\ z + r \end{pmatrix}\right\} \qquad (\operatorname{mod} p)$$

Si $i = j = r$, l'autotopie A devient un automorphisme par le groupe cyclique.

Supposons $i, j, r \neq 0$ et faisons subir à $G(.)$ l'isotopie

$$T\left\{\xi = \binom{x}{xi'}\ ;\ \eta = \binom{y}{yj'}\ ;\ \zeta = \binom{z}{zr'}\right\} \qquad (\mathrm{mod}\ p)$$

où i', j', r' sont les associés de i, j, r: $ii' \equiv jj' \equiv rr' \equiv 1 \pmod p$. Soit $\Gamma\ (*)$ l'image de G par T: $\Gamma = G_T$.

Dans l'autotopie A, l'image de x est $x + i$, dans Γ l'image de xi' est $(x + i)$ $i' = xi' + 1$. Donc Γ est invariant par l'autotopie

$$\left\{\binom{x}{x+1}, \binom{y}{y+1}, \binom{z}{z+1}\right\}$$

c'est-à-dire qu'il est automorphe par le groupe cyclique.

(iii) Supposons une dès trois quantitités i, j, r, et une seule, par exemple j, nulle. $G_{i,0}$ est toujours automorphe par:

$$A\left\{\xi = \binom{x}{x+i}, \eta = \binom{y}{y}, \zeta = \binom{z}{z+r}\right\}.$$

Faisons l'isotopie:

$$T_1\left\{\xi_1 = \binom{x}{xi'}, \eta_1 = \binom{y}{y}, \zeta_1 = \binom{z}{zr'}\right\};$$

nous obtenons un quasigroupe $\Gamma_1(*) = GT_1$, qui est cette fois invariant par l'autotopie

$$\left\{\binom{x}{x+1}, \binom{y}{y}, \binom{z}{z+1}\right\}$$

car la relation $x.y = z \Rightarrow (x + i).y = z + r$, devient par T_1 $xi'*y = zr' \Rightarrow (xi' + 1)*y = zr' + 1$. Donc $x*y = z \Rightarrow (x + 1)*y = z + 1$.

Exemple. Γ_1: $x*y \equiv x + 3y + 3 \pmod 5$.

Or si $0*y = a$, on aura: $x*y = a + x$, donc si l'on applique à Γ_1 l'isotopie.

$$T_2\left\{\binom{x}{x}, \binom{y}{a}, \binom{z}{z}\right\},$$

on aboutira à un quasigroupe $\Gamma\ (\wedge)$ tel que: $x \wedge a = x + a$, ce qui est le groupe cyclique d'ordre p. Or

$$\binom{y}{a} \to \binom{y}{0*y} = \binom{y}{0.y}\zeta'$$

donc enfin

$$\Gamma = GT_1T_2,\ T_1T_2 = T'\!:\!\left\{\xi' = \binom{x}{xi'}, \eta' = \binom{y}{0.y}\zeta', \zeta' = \binom{z}{zr'}\right\}$$

Si l'élément nul est i à la place de j on arrive à une conclusion symétrique.

Exemple. $n = 25$; $C_0 = [0, 5, 10, 15, 20]$; $C_2 = [2, 7, 12, 17, 22]$; $C_2 \times 0 = [14, 24, 9, 19, 4]$; $C_2 \times 5 = [24, 9, 19, 4, 14]$; $C_2 \times 10 = [9, 19, 4, 14, 24]$; $C_2 \times 15 = [4, 14, 24, 9, 19]$; $C_2 \times 20 = [19, 4, 14, 24, 9]$; $i = 2, j = 0, r = 4$. Le quasigroupe $G_{2,0}(.)$ est $G = [0, 1, 2, 3, 4]$; $G.0 = [2, 4, 1, 3, 0]$; $G.1 = [4, 1, 3, 0, 2]$; $G.2 = [1, 3, 0, 2, 4]$; $G.3 = [0, 2, 4, 1, 3]$; $G.4 = [3, 0, 2, 4, 1]$;

$$T'\left\{\xi' = \begin{pmatrix} 0\,1\,2\,3\,4 \\ 0\,3\,1\,4\,2 \end{pmatrix}, \eta' = \begin{pmatrix} 0\,1\,2\,3\,4 \\ 3\,1\,4\,0\,2 \end{pmatrix}, \zeta' = \begin{pmatrix} 0\,1\,2\,3\,4 \\ 0\,4\,3\,2\,1 \end{pmatrix}\right\}$$

$\Gamma = G_{T'}$ est le groupe cyclique additif de $Z/5$.

(iv) Si $i, j \neq 0$ et $r = 0$, $G(.)$ est invariant par l'autotopie

$$A\left\{\xi = \begin{pmatrix} x \\ x+i \end{pmatrix}, \eta = \begin{pmatrix} y \\ y+j \end{pmatrix}, \zeta = \begin{pmatrix} z \\ z \end{pmatrix}\right\}.$$

Faisons l'isotopie

$$R_1\left\{\xi_1 = \begin{pmatrix} x \\ xi' \end{pmatrix}, \eta_1 = \begin{pmatrix} y \\ yj' \end{pmatrix}, \zeta_1 = \begin{pmatrix} z \\ z \end{pmatrix}\right\},$$

nous obtenons le quasigroupe $\Omega\ (*) = G_{R1}$. La condition

$$(x + i)\cdot(y + j) = x.y$$

devient par R_1 $(x + 1)*(y + 1) = x*$y.

Si l'on fait subir à Ω une nouvelle isotopie

$$R_2\left\{\xi_2 = 1, \eta_2 = \begin{pmatrix} y \\ -y \end{pmatrix}, \zeta_2 = \begin{pmatrix} z*0 \\ z \end{pmatrix}\right\}$$

on obtiendra le groupe cyclique $\Gamma : x \wedge y = x + y \pmod{p}$. Finalement, $G(.)$, multiplié par l'isotopie $R_1R_2 = T''$

$$T''\left\{\xi'' = \begin{pmatrix} x \\ xi' \end{pmatrix}, \eta'' = \begin{pmatrix} y \\ -yj' \end{pmatrix}, \zeta'' = \begin{pmatrix} z.0 \\ zi' \end{pmatrix}\right\}$$

devient le groupe cyclique.

Exemple. Soit le quasigroupe "endo" d'ordre 25, défini par $[0, 5, 10, 15, 20] \times 1 = [16, 1, 6, 21, 11]$; $1 \times [0, 5, 10, 15, 20] = [7, 17, 22, 12, 2]$; $[1, 6, 11, 16, 21] \times 1 = [9, 14, 19, 24, 4]$; $[2, 7, 12, 17, 22] \times 1 = [10, 0, 15, 5, 20]$; $[3, 8, 13, 18, 23] \times 1 = [18, 8, 23, 13, 3]$, et $[4, 9, 14, 19, 24] \times 1 = [17, 22, 2, 7, 12]$. On a $G_{4,2} = [0, 1, 2, 3, 4]$; $G.0. = [4, 2, 0, 3, 1]$; $G.1 = [3, 1, 4, 2, 0]$; $G.2 = [2, 0, 3, 1, 4]$; $G.3 = [1, 4, 2, 0, 3]$; $G.4 = [0, 3, 1, 4, 2]$. On en tire $\Omega = [0, 1, 2, 3, 4]$; $\Omega *0 = [4, 1, 3, 0, 2]$; $\Omega *1 = [2, 4, 1, 3, 0,]$; $\Omega *2 = [0, 2, 4, 1, 3]$; $\Omega *3 = [3, 0, 2, 4, 1]$; $\Omega *4 = [1, 3, 0, 2, 4]$. Et Ω_{R2} est le groupe cyclique d'ordre 5.

(v) Si enfin deux nombres sont nuls, le troisième l'est aussi et $i = j = r = 0$. $G_{0,0}$ coïncide avec l' "endo" d'ordre p isomorphe à D par $(x \to x/p^{\alpha-1})$.

Ce théorème permet de construire par récurrence les "endo" d'ordre p^2, p^3, . . . à partir des "endo" d'ordre p. La construction des "endo" d'ordre premier résulte de la proposition suivante.

27. Cas où *n* est premier. *Pour qu'un groupoïde $G(\times)$, d'ordre premier, p, automorphe par le groupe géométrique (et par conséquent "endo") soit un quasigroupe, il faut et il suffit que les fonctions*

$$x \times 1 = f(x) \text{ et } x'f(x),\ xx' \equiv 1 \qquad (\text{mod } p),$$

définissent deux substitutions $(x \rightarrow f(x))$et $(x \rightarrow x'f(x))$, où la valeur de $x'f(x)$ correspondant à $x = 0$ est conventionnellement prise égale à celui des nombres $0, 1, 2, \ldots, p-1$ *qui ne figure pas parmi les $p-1$ valeurs de $x'f(x)$. Le quasigroupe est alors entièrement déterminé par la fonction $f(x)$.*

Exemple.

$$\begin{aligned} x &= [0, 1, 2, 3, 4, 5, 6] \qquad (\text{mod } 7).\\ f(x) &= (3, 5, 2, 4, 1, 6, 0]\\ x' &= [-, 1, 4, 5, 2, 3, 6]\\ x'f(x) &= [-, 5, 1, 6, 2, 4, 0];\ x'f\ (0) = 3 \end{aligned}$$

(*cf.* N° 4, Exemple II).

La démonstration directe de cette proposition est facile, mais lourde; la propriété résulte d'ailleurs immédiatement, par le N° 4, de l'énoncé corrélatif dans le cas des groupoïdes automorphes par le groupe cyclique (**12**, N° 6).

Si un quasigroupe d'ordre premier est monogène, il est clair que son automorphe se réduit au seul groupe géométrique. Mais cette condition n'est pas nécessaire et l'on peut construire des quasigroupes d'ordre premier, possédant des diviseurs, et dont le groupe d'automorphisme se réduise néanmoins au groupe géométrique. Ainsi le quasigroupe "endo" du 19$^{\text{ème}}$ ordre défini par la substitution

$$(x \rightarrow x \times 1) = (0, 18, 13, 17, 2, 10, 6, 16, 15, 3, 5, 12, 9)(1, 7)(4, 8, 14)(11),$$

écrite sous forme de produit de cycles, admet six diviseurs monogènes isomorphes du 3$^{\text{ème}}$ ordre: [1, 7, 11] et ses produits par 2, 4, 5, 8 et 10: [2, 14, 3], [4, 9, 6], [5, 16, 17], [8, 18, 12], et [10, 13, 15]. Son automorphe se réduit aux 18 transformations $(x \rightarrow mx)$, $(m = [1, 2, 3, \ldots, 18])$. En effet, parmi les séries des multiplications à droite (**13**, N° 7-8), une seule est d'ordre 18; elle est formée par les puissances de 2 (mod 19). Comme elle doit se projeter sur elle-même par tout automorphisme, A_Q est d'ordre 18.

Citations

1. A. A. Albert, *Non-associative algebras I.* Ann. Math., *43* (1942), 696.
2. ———, *Quasigroups*, Trans. Amer. Math. Soc., *54* (1943), 510.
3. C. Burstin, W. Mayer, *Distributive Gruppen von endlicher Ordnung*, J. reine ang. Math., *160* (1929), 111–130.
4. A. Cauchy, *Exercices d'Analyse et de Phys. Math.* III (Paris, 1844).
5. L. E. Dickson, *History of the Theory of Numbers*, II (New York, 1952).
6. R. A. Good, *On the theory of clusters*, Trans. Amer. Math. Soc., *63* (1948), 482–513.
7. G. H. Hardy and E. M. Wright, *Number theory* (Oxford, 2$^{\text{ème}}$ édit., 1953).
8. B. A. Hausmann, O. Ore, *Theory of quasigroups*, Amer. J. Math., *59* (1937), 983–1004.
9. D. C. Murdoch, *Structure of Abelian quasigroups*, Trans. Amer. Math. Soc., *49* (1941), 395.
10. L. J. Paige, *Neofields*, Duke Math. J., *16* (1949), 39–60.
11. A. Sade, *Quasigroupes* (Marseille, 1950).
12. ———, *Groupoïdes automorphes par le groupe cyclique*, Can. J. Math., *9* (1957), 321–335.
13. ———, *Quelques remarques sur l'isomorphisme et l'automorphisme des quasigroupes*. Abh. Math. Sem. Univ. Hamburg (1958).
14. O. Schmidt. Math. Z., *29* (1929), 34–41.
15. G. Scorza. *Gruppi Astratti* (Roma, 1942).
16. M. F. Smiley. *Application of a radical of Brown, McCoy to non-associative rings*, Amer. J. Math., *72* (1950), 93–100.
17. B. M. Stewart, *Theory of Numbers.*
18. A. Suschkewitsch, *Ueber die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit*, Math. Ann., *99* (1928), 33.
19. M. Takasaki, *Abstraction of symmetric transformations*, Tôhoku Math. J., *49* (1943), 145–207.

Lycee Perier, Marseille

ntal
ract
life,
958.

n in
pon
tics
cts.

y of
ised
re-
onal

tion
rom
958.

the
rent
rical

ply.
958.

and

pec